ARL-228
Part II

PASCUAL JORDAN

FINAL REPORT II, 1961

WITH CONTRIBUTIONS OF W. KUNDT
J. EHLERS

PROBLEMS OF GRAVITATION

Containing two parts: Empirical aspects of Dirac's hypothesis. – Progress in the mathematical problems of Einstein's theory of gravitation

PASCUAL JORDAN

FINAL REPORT II, 1961

WITH CONTRIBUTIONS OF W. KUNDT,

J. EHLERS

PROBLEMS OF GRAVITATION

Containing two parts: Empirical aspects of Dirac's hypothesis. - Progress in the mathematical problems of Einstein's theory of gravitation.

The research reported in this document has been supported in part by the

AERONAUTICAL RESEARCH LABORATORY

of the OFFICE OF AEROSPACE RESEARCH, UNITED STATES AIR FORCE

through its European Office

About this Report

Though this report in its whole extent deals with problems of GRAVITATION, its first half and its second half are quite different in nature. In the first half empirical results of several branches of natural sciences are discussed, in relation to a fundamental problem of gravitation. But in the second half mathematical developments are reported, belonging to the consequences deduced from EINSTEIN's theory of gravitation and general relativity. In order to show to the reader how I came to begin research - together with my pupils and collaborators - in these widely different directions, and with an intention to look at these different endeavours as parts of an undivided program, I should like to make a few remarks about the connections in my scientific work during the last fifteen years.

My interest in general relativity arose when I became acquainted with Dirac's cosmological speculations, which lead him 1937 to his hypothesis about cosmologically diminishing gravitation. I have been then nearly the only one among physicists who was really fascinated by Dirac's idea and convinced by his arguments, and I tried to think more about them. I saw connections of these ideas with the interesting developments of Kaluza, O. Klein, Veblen, W. Pauli about five dimensional or "projective" relativity (these connections have been detected and discussed also by Einstein and P.G. Bergmann), and I felt myself necessitated to study more thoroughly general relativity, a branch of modern physics to which before these events I had devoted only cursory attention, having been busied with the fascinating problems of quantum mechanics and quantum electrodynamics or quantum field theory - which I had to lay aside now in order to study gravitation.

The first period of my new endeavour since 1944 concerned the task to generalise Einstein's theory of gravitation in such a manner as to include in the new "generalised theory of gravitation" possibilities of the kind suggested by Dirac. Much has been done in this direction not only by myself but also by my friends G. Ludwig, Cl. Müller, O. Heckmann, W. Pauli, W. Fricke, H. Greßmann, H. König, E. Schücking, J. Ehlers, W. Kundt, K. Just and independently by Y. Thiry. Important contributions

to these matters have been given by M. Fierz, A. Lichnerowicz, and then by R. H. Dicke and D. Brill.

But in the course of time it had become clear to me that those mathematical problems of the "generalised theory of gravitation" which seemed to me to be chiefly interesting could not be studied successfully without having at first studied thoroughly the whole field of mathematical problems in the realm of Einstein's theory - with a real "constant" of gravitation. Therefore since several years I encouraged my young friends to do research chiefly in this direction - the second half of this report presents results in this field.

I prefered to do so also because it became a little difficult for me to maintain my hope that it might be really useful to do work concerning the "generalised theory of gravitation" - astronomers seemed to remain undividedly inclined to believe that this theory (and the underlying hypothesis of Dirac) might have nothing to do with reality. Though I myself remained convinced of the correctness of this theory, it seemed to me during some years not to be justified to induce my pupils to do more about a theory of doubtful value.

Therefore I became more and more interested in the question whether there are empirical facts which might be regarded as empirical evidence in favour of Dirac's hypothesis and of my own theory, based upon this hypothesis and developing it. This task grew to a seperate study winning more and more interest - and what is reported here in the first half of this report, seems to me now to be a definitive statement showing Dirac's hypothesis to formulate one of the fundamental truths about the laws of Nature.

Hamburg
Okt 60.

P. Jordan

Technical Remark.

In Chapter I,II the author had to mention many details of sciences not belonging to his own branch - such as geology, palaeoclimatology and so on. Writing down this report he felt that many technical words out of the scientific vocabulary of these sciences are not known to him with certainty in their English version. Having not the ambiticn to deliver a work of philological perfection, he decided not to go in each single case to the libraries of other institutes, but to express himself in a suitable manner so that the kind reader without doubt will understand him, though missing some customary designations.

PART I. DISCUSSION OF EMPIRICAL FACTS GIVING SUPPORT TO DIRAC'S HYPOTHESIS OF A VARIABLE "CONSTANT" OF GRAVITATION.

§ 1. Introduction. Dirac 1937 by considerations about which we shall speak only later came to the conclusion that what we use to call the gravitational constant may be in reality a variable. The force of gravitation between two masses of one gramme each, and each one of spherical symmetry, in the case that their centres are one centimeter apart, has the well known value

$$f = 6{,}685 . 10^{-8} \text{ dyn.cm}^2.\text{g}^{-2}. \tag{1}$$

In the relativistic theory of gravitation, as developed by Einstein and other authors, we often use the notation

$$\varkappa = \frac{8\pi^2}{c^2} f . \tag{2}$$

The meaning of Dirac's hypothesis is now this: In the course of geological or cosmological times this f performed a process of slow monotonuous diminution. If f_o is the value of f at present - equal to what we called f in (1) - then the diminution Δf of f during one year may have been during the last 10^9 years of the order of

$$\Delta f = 10^{-9} f_o . \tag{3}$$

But Dirac's quite speculative considerations gave no exact value of this amount.

Looking seriously at this idea of Dirac's, and knowing the fundamental concepts of relativity, one must conclude that f or $\varkappa$, if variable in time, must also be variable in space; it must be what the theory of relativity calls a scalar field variable. There arises then the problem to formulate new field equations of the theory of general relativity, and to discuss the theoretical consequences. The author began to study this topic 1944. Many authors participated in this mathematical and physical study; I mention here only the names of Thiry, Pauli, C. Müller, Ludwig, Heckmann, Fricke, Ehlers, Schücking, Kundt, Fierz, Greßmann, Lichnerowicz, Dicke, Brans, Brill, Leutwyler, Just. Though it might seem that there are scarcely sufficient criteria to allow from a purely theoretical point of view, precise (hypothetical) statements about these new field equations, the situation in reality is a different one. From Einstein's admirable theory of gravitation we know the possibility that theoretical considerations possessing a high convincing power may lead us in a field of problems where empirical evidence is still extremely poor: only now, with the help of satellites, we can hope to get more and deeper information about finer details in the laws of gravitation; and nobody among physicists will doubt that this future experimental evidence will in the first line give - as the Mossbauer effect - confirmation of the theoretical predictions drawn from Einstein's theory.

Only the numerous attempts - starting from Weyl's famous and mathematically brilliant, but surely physically wrong theory - to find what has been called a unitary theory of gravitation and electromagnetism (or even also fields of meson forces) lead us into speculations without limits and without any hope to get real physical predictions about possible experiments of the future. This whole endeavour, put forward by so many leading mathematicians and physicists, must be said to be probably entirely valueless, in contrast to Einstein's original theory of gravitation (including also Maxwell fields, but without the ambition to give a "geometric interpretation" of these or a "unitary theory") which must be acknowledged as a probably correct description of deep physical laws not yet explored experimentally in a sufficient manner.

Introducing now, according to Dirac's idea, an additional scalar field variable into the theory we naturally get a modification - or a generalisation - of Einstein's classical theory of gravitation. Though again we have to do here steps of a decidedly hypothetical character, we find again such a situation that there can be scarcely any doubt about the possibility of a simple and convincing generalisation of Einstein's theory in the direction searched for.

Let us write down - restricting us here to the vacuum case - the fundamental variational principle of the combined Einstein-Maxwell theory. In usual notation it is:

$$\delta \int \left(\mathcal{G} - \frac{\varkappa}{2c^2} F_{kl} F^{kl} \right) d\tau = 0 \tag{4}$$

with

$$d\tau = \sqrt{-g}\, dx^{(0)} \dots dx^{(3)}. \tag{4'}$$

To come now to the modified new variational principle of the generalised theory of gravitation ("erweiterte Gravitationstheorie") we write [1)]

$$(5) \qquad \delta\int\left(\frac{\mathcal{G}}{\varkappa} - \zeta\,\frac{\varkappa^{|\nu}\varkappa_{|\nu}}{\varkappa^{3}} - \frac{1}{2c^{2}}\,F_{kl}\,F^{kl}\right)d\tau = 0;$$

$$(5') \qquad \varkappa^{|\nu}\varkappa_{|\nu} = g^{kl}\,\frac{\partial\varkappa}{\partial x^{k}}\,\frac{\partial\varkappa}{\partial x^{l}}.$$

Here the dimensionless constant ζ is the only one unknown parameter of the theory, to be evaluated not by theoretical speculation but from empirical facts. There are reasons to believe that ζ should be great in comparision with 1, perhaps of the order 10^2; but according to Dicke the empirical evidence (from the perihelion of Mercury) is perhaps not so sharp as to make a value of the lower order 10 impossible.

The great bulk of mathematical and physical discussion and research in connection with (5) may be excluded here from consideration, because the authors and his collaborator's extensive studies about this topic do not belong to the period covered by this contract. New mathematical research in this direction has become now of actual urgency, but is only in its beginning.

Y. Thiry, who too studied the generalisation of Einstein's theory of gravitation, necessary if one tries to study the theoretical consequences of Dirac's idea, independently of the author, came to similar developments. R.H. Dicke, starting from a study of empirically detectable consequences of Dirac's

---

1) Discussion of the present status of affairs: P. Jordan, Zeitschr. Physik 157, 112 (1959).

hypothesis (5), developed together with Brans also the formulation of the relativistic field equations for the theory with variable $\varkappa$, in connection with a discussion of Mach's principle [1]. A very clear and instructive discussion of the theoretical foundations of the theory, showing thoroughly also the relations between the endeavours of the different participating authors, has been performed by D. Brill [2].

In this report not much more will be said about the theoretical side of the problem. I prefer to discuss the empirical aspect, summarising previous discussions and going further some steps.

Experimental tests apt to decide about Dirac's hypothesis are closely related to the geophysical questions influenced by the hypothesis. Their discussion therefore may be included into the general discussion of those facts concerning our Earth which are of interest in this connection. Our discussion then consists of two chapters: The second one will treat facts of astronomy. The first one will deal with the earth. Some remarks about other planets and about the moon may be included in the first chapter.

---

1) C.H. Brans, R.H. Dicke, Mach's principle and a relativistic theory of gravitation. 1961,- C.H. Brans, Mach's principle and a varying gravitational constant. Dissertation Princeton 1961.

2) Review of Jordan's extended theory of gravitation. Varenna 1961. International School of Physics "Enrico Fermi".

CHAPTER I. CONSEQUENCES OF DIRAC'S HYPOTHESIS CONCERNING PHYSICS AND HISTORY OF THE EARTH.

§ 2. Possible experimental tests. There exists perhaps some hope to measure in a near future f with such an accuracy that during 10 or 20 years the predicted small diminution could be detected. I learned from R. Vieweg that the measurement of gravity at the surface of the earth will come soon to the precision of at least 6 decimals; and by suitable arrangement one may perhaps win still more 3 decimals for comparision measurements. Gravimetric measurements allow already to discern periodical variations of gravity corresponding to the 9. decimal of f.[1] The best present hopes in this direction seem to me to arise from the work of Dicke and collaborators in Princeton who are endeavouring to attack the problem of precision measurement of gravity in several ways.

Diminution of f must lead to a slow expansion of the earth, influencing as well as f itself the gravitational acceleration on the surface of the earth. It is a favourable fact that both effects - the direct effect of the hypothetical change of f, and our increasing distance from the centre of the earth - go into the same direction: The amounts of both effects may perhaps be of the same order of relative magnitude 10^{-9} or 10^{-10}

---

1) R. Tomaschek, Handbuch d. Physik, XLVIII, page 775. (Bln - Göttingen - Heidelberg 1957).

per year. But I myself, being now nearly in my sixtieth year, have not much hope that a clear decision of this great problem of natural science by direct measurement will come out already during my lifetime.

Variation of the radius of the earth must also give variations in its rotational period - the law of conservation of rotational momentum remains valid also in the generalised theory of gravitation. Therefore there might seem to be chances to detect something about the hypothesis by supervision of the rotation of the earth, modern methods with SiO_2 clocks, NH_3 clocks, Cs clocks allowing highly precise measurement. Indeed variations of the rotational velocity of the earth of the magnitude to be expected from the hypothesis would fall into the order of magnitude of the really known empirical deviations from constant rotation. [1] But these deviations partly are caused by effects which are well understandable in a qualitative manner, but not easy to compute with the necessary accuracy. We have a periodical variation of the rotation in consequence of the fact that about 10^{-8} of the rotational momentum of the earth is contained in the wind system and in the tidal movements. (Both these parts are of the same order of magnitude). Apart from these periodical variations there are aperiodical ones amounting to only about 1% of the periodical ones, - they change the duration of the day in the order of 1 msec pro 100 years. It is not improbable the slow variations of oceanic currents and also the melting process of polar ice

1) H. Spencer-Jones, Handb. Physik XVLII, page 1. (1956).

are partly responsible for these aperiodical changes of rotation, but exact computations are scarcely possible.

Friction of tidal movement causes a certain transfer of rotational momentum from the earth to the moon. This effect has been discussed quantitatively with the following result: Astronomically a slow increase of the rotational momentum of the moon, relative to the earth, is known. Partly this is caused by perturbations by other planets; but in the same order of magnitude also tidal friction is partly responsible for this astronomical effect. But naturally the tidal friction cannot be computed theoretically with high accuracy; therefore we can scarcely hope to be able so learn something about Dirac's hypothesis from the rotation of the earth.

The chances are even still worse: Some specialists doubt that the increase of rotational momentum of the moon really gives the corresponding decrease in the rotation of the earth. For there are reasons to presume that the tidal movement of the atmosphere allows a compensation for this loss of rotational momentum of the earth, by a resonance effect discussed already by Kelvin. (The rotational momentum of the earth relative to the sun makes such a compensation possible without violation of the conservation law). At last the secular variation of the earth's magnetic field seems to show that the rotation of inner and outer parts of the earth is not exactly synchronised but shows extremely slight differences. All these facts together make it an hopeless task to get information about the surmised variation of $\varkappa$ from the rotation of the earth.

Conservation of rotational momentum in a two body problem

with gravitational force has the consequence that in the case of variation of $\varkappa$ the average radius of the orbit varies proportional to $\varkappa^{-1}$, and the period of revolution proportional to $\varkappa^{-2}$. The corresponding variation in the length of a year cannot be measured astronomically. But in the case of the moon and its orbit around the earth the situation is more favorable. According to Dirac's hypothesis the moon should show a "secular acceleration" of about - 0,5" . In fact it shows a secular acceleration of +10", and this is the effect of two combined causes, giving each one about one half of the effect: Planetary perturbations, computed already with old methods by Laplace, could today be determined quite exactly by use of modern electronic computing devices. But the other cause is again tidal friction - and there is scarcely any chance that this could be computed theoretically with a precision of 5% or 1% als would be necessary in order to allow recognition of the effect of diminishing $\varkappa$.

Now here is one of the great chances of modern satellite technic to help us to look into the fundamental laws of Nature. For satellites moving around the earth naturally do not undergo any influence of tidal friction - this friction being a function of the quotient of the height of tidal waves to the depth of the ocean. (Practically the wast shallow Bering Sea contributes the chief part of the tidal friction of the moon).

The technical difficulty of weak atmospheric friction of a satellite can be removed by an idea discussed by

American physicists. I learned about it by the kindness of M. Schwarzschild. In principle a satellite of the needed type would consist of two parts: The central body, and the mantle surrounding it but not connected with it. The central body moves without any frictional contact with the atmosphere. The mantle, loosing slowly by friction some of its momentum, contains a) instruments to measure any deviation of the central body from its normal position in the mantle; b) a suitable rocket device in order to accelerate it automati= cally in the needed manner in order to restore the normal spacial relation between mantle and body.

§ 3. The reality of earth expansion. In several ways any considerable change concerning $\varkappa$ must have strong influence upon the structure of the earth and the processes going on at its surface. If really $\varkappa$ did - during geolo= gical times - perform a diminution of such an amount as Dirac's hypothesis tends to make us believe, then this must have lead to consequences detectable in several different branches of the sciences endeavouring to understand our earth and the phenomena presented by it. We try to discuss this matter in the following paragraphs to its full extent, though it will then be necessary to look into a series of problems which usually are thought to lie far beyond the limits of physical interests.

The aim of this whole discussion will be to clear whether there would arise from Dirac's hypothesis consequences which seem to be in contrast with empirical facts - possibilities of this kind could be numerous, and only by a thorough study

we can come to see whether this hypothesis may be discussed by the physicists seriously without challenging vehement criticism of other scientists. In this connection we have to mention various problems which have been discussed by many authors without gaining generally acknowledged results. If we aim to show how these controversial matters seem to be thought of as soon as we are following Dirac's idea, primarily we tend to show that Dirac's hypothesis cannot be said to be in contradiction with well established facts. In many details, to those who are interested exclusively in these details, it may seem that our interpretation, based on Dirac's hypothesis, is not a necessary one because other theories too are perhaps apt to give an explanation for just this or that detail. But our point is to examine whether there are hindrances to regard the hypothesis as a possible one, in harmony with what we know about our earth.

Therefore we have here the new and perhaps a little surprising situation that facts and discussions of geology, plaeoclimatology, palaeomagnetism, Volcanism, oceanography and so on win some meaning about a problem interesting primarily the physicist.

As already said above the scalar $\varkappa$, if variable in the course of time, must be variable also spacially; and there exists now at least one attempt to look for an effect suited as a test about this possibility. In the centrally symmetric static gravitational field of a mass point or a star $\varkappa$ is a function of the distance from the centre; the mathematical problem to determine this function is solved since more than

ten years. Now Dicke proposed the following idea: Owing to the excentricity of its orbit the earth comes in the course of the year into districts where $\varkappa$ is a little greater or smaller. This must cause a slight periodical variation in gravitation at the surface; one can surmise that per= haps this leads to a seasonal variation in the frequency of earthquakes.

Morgan, Stoner and Dicke[1] made a study of this question. Indeed the harmonic analyse of the empirical frequencies shows a variation which is in accord with Dickes's expectations.

All other empirical facts to be discussed in this chapter are related to the assumed dependency of $\varkappa$ from time. Two main consequences are to be drawn: [2] 1) The earth must have performed, as mentioned already, an expansion.[3]
2) The "solar constant", measuring the intensity of solar radiation at the earth, must have diminished strongly in the course of geological times. The second point will be discussed later; at first we have to do here with the expansion of the earth, and related topics. When 1952 with kind permission of my friend Joel Fisher I discussed in a book of mine his idea that such an expansion should have taken place, this seemed to be quite a revolutionary idea, in sharp contrast to current theories of geologists and geophysicists believing in a contraction of the earth. Today the situation is quite another one: The opinions of leading specialists in this field are converging now to acknowledge

---

1) W.J. Morgan, T.O. Stoner, R.H. Dicke, Periodicity of earthquakes and the invariance of the gravitational constant. princeton 1961

expansion of the earth as an empirically stated fact; differences of opiniom remain only concerning the amount of this expansion.

The decisive fact - revealed by Ewing, Heezen, Tharp (8) - is the existence of a great system of deep rifts in the surface of the earth. Heezen (4) has given a very instructive description of this system and analysed its significance. Also J.T. Wilson (7) acknowledged the expansion of the earth as a fact proven by this system of rifts. I shall not go here into details, but give only a few indications about this phenomenon. Fig.1 shows in the first line those parts of the system of rifts which are lying in the oceans. Several smaller parts of these rifts (extending to lengths of hundreds of km) have been already well known since some time. The depths of these rifts, as one knows, goes to the amount of about 10 km below the surface of the oceans; their widths are sometimes only 2 km. The existence of this whole system has been discovered by Ewing, Heezen, Tharp in connection with their oceanographic research; oceanographic results of other sources - especially also the nuclear submarines Nautilus and Skate, gave confirmations and extensions.

Some parts of these rifts are lying on the continents - especially the famous east african rifts, where especially in the case of the Great Dike in Southern Rhodesia in the length of 500 km a rift of 3 to 5 km is filled by magma from below. This phenomenon did already since some time cause difficulties for the usual theory of a contracting earth. These difficulties win overwhelming power now after the discovery that the Great Dike is only a small part of a system enveloping the whole earth. There are probably on the continents still many other rifts belonging to this system, but in

most cases detectable only by thorough geological investigation. An example of these seems to be a smaller system of rifts studied by Stille (9); fig.2 shows that these rifts are lying in the continuation of one of the parts of the system fig.1.

One of the great arms of these rifts is going over Iceland; a photo made from a plane reproduced in Heezen's mentioned report shows quite visibly the nature of these rifts as consequence of a pull tearing the adjacent parts of the surface from each other. The rift cuts across Iceland through the depression known as the Central Icelandic Graben. All the recent volcanic activity of the island is limited to this valley, and almost all of its earthquakes originate there. It has been measured that the Graben is growing wider at the rate of 3,5 m per 1000 years for every km of its width.

The rift cutting across iceland is ~~lying~~ extended along the atlantic ridge ("atlantische Schwelle") showing this to be the result of vively volcanic activity accompanying (or caused by) the tearing process in the rift.

The continental parts of this system of rifts probably are in many cases old ones, showing today scarcely any activity. But the oceanic parts show themselves to be mostly in recent activity: 1) Recent earthquake epicenters are accumulated along these rifts. 2) An intensified flux of heat out of the depth is coming up in these rifts. 3) Gravimetric anomalies are known since former years to exist along parts of these rifts - they prove that we are observing here young structures in relatively rapid transformation.

Expansion of the earth - contrary to the former idea of contraction - now seems to be a well established fact. If the cause of this expansion is indeed a diminution of $\varkappa$, according to Dirac's hypothesis, then this is not a special feature of the earth - but the earth gives only an example of a universal phenomenon. There are only two other celestial bodies which could allow us any detection of such an expansion, if real: Our moon and the planet Mars.

Concerning the moon it is well known that there exist certain rifts ("Rillen") which are interpreted by the specialists indeed as proof of a slight expansion. (The interested authors believed that temperature increase caused by radioactivity might explain this expansion). In the case of the earth the rifts in the oceans must, as already said, be interpreted as structures of short duration, so that they cannot give direct information about the total amount of expansion during long times. But the rifts in the surface of the moon scarcely can have been wiped out again partially since the time of formation of the great craters and the mares. (The mentioned rifts are younger formations than most of the craters). Therefore the sum of the width's of these rifts may measure the total expansion of the moon since that time - it is in harmony with Dirac's idea that in the case of the moon the total expansion must be very small in comparision with the earth.

Concerning **Mars**, the famous "canals" probably are symptoms of expansion. As one knows, these canals in former times have been the object of wild, phantastic speculation about inhabitants of Mars and their highly developed technical activities. Therefore during long times it seemed to be appaesing to believe that all these canals might only be optical illusions. But today again the best informed specialists are inclined to think that these "canals", though not corresponding to this name, are real phenomena; and it seems to be rational to surmise that they indeed indicate expansion also of Mars. Surely the coming developments of space research will enable us to look clear and to get undoubtable information about this problem.

Speaking about the moon we may also mention that **H. Krause** [1] from a careful study of formations on the moon came to the conclusion that there must have been in former times a real atmosphere. At present there are, as one knows, ~~already~~ only small traces of an atmosphere - causing however that in greater altitudes, (about 80 km), comparable which those which in the case of the earth are the scene of shooting stars, the lunar atmosphere (loosing density with increasing altitude much more slowly then our own atmosphere) has the same density as the atmosphere of the earth in the corresponding altitude, so that also from the moon shooting stars are to be seen.

Krause recognized on the moon the existence - in crowds -

---

1) Die Sterne, 28, 199(1952)

of formations which he thinks must be interpreted as wind ero= sions and as aeolic sediments. If his interpretation is the correct one, then it would be scarcely possible to deny that the moon in former times has been capable - in consequence of a greater $\varkappa$ - to bind an atmosphere.

A similar result comes from known facts about Saturn's moon Titan, which has today an atmosphere, but which would - according to Kuiper - be unable to maintain this atmo= sphere if its very low surface temperature would be higher. Kuiper drew from this fact the conclusion that Titan during its formation must have had about the same low temperature as now - a conclusion which seems to me to be quite far from any probability.

§ 4. The formation of the ocean bassins. Though it seems that expansion of the earth - contrary to what geologists believed during more than a century - must now be considered as a proven fact, it remains a controversial question whether the amount of expansion might be great enough in order to allow also the phenomenon of ocean basins to be interpreted as a result of expansion. We prefer to think that this question cannot yet be said to be answered definitely. But strong arguments have been put forward in favour of this interpretation which has been given independently by three different authors.

1) J. Fisher came to this interpretation of ocean basins in the course of a discussion (between him and the

author) of Dirac's hypothesis. His idea has been published with his kind permission 1952 in a book of the author.(1). Later on (1955) I tried to give more details about this idea in a second edition of the book.

2) L. Egyed 1957 gave the same interpretation - ocean basins as a result of earth expansion - at the basis of empirical facts of geology and geophysics. At that time, not yet acquainted with Dirac's hypothesis, he suggested in a speculative manner other physical causes of this expansion. But later on he acknowledged (in a letter to the author) Dirac's hypothesis as a quite convincing explanation of this expansion, which he is inclined to think of as an empirical fact.

3) B.C. Heezen 1960 came to the same conclusion - expansion as cause of ocean basins - and contributed very convincing new material to prove this conclusion as the correct one. Heezen too discussed the whole matter strictly from an empirical point of view, without deciding which physical cause of the expansion might be probable.

The author of this report has the impression that not only new, recently discovered facts, but also former, well known results concerning the relations between continents and ocean basins give strong support to the conclusion of strong expansion, though some authors, as Dicke (5) and Wilson, prefer to acknowledge only some weak expansion as a proven

fact. It may therefore be useful to discuss in the following also the fundamental facts about continents and oceans: They are well known, but their significance has not yet generally been seen clearly enough.

Fact one is the existence of two different levels of altitudes on the surface of the earth. This well known phenomenon is shown here in fig 2 and fig 3 which do not necessitate details of comment. The water of the oceans by chance is sufficient to fill the areas of the deeper level just according to the altitude difference of the two levels, making the areas of the deeper level to ocean basins, but covering only bordering districts of the continental blocks, the shelves, with shallow water.

We add fig 4 in order to emphazise that the Pazific is, with respect to hypsography, not essentially different from the other oceans. This is important because a certain quite phantastic theory has been put forward according to which the Pacific would be a phenomenon completely different in nature and origin from the other oceans. Following the idea of Darwin that the moon might have originated by a process of separation from the earth, some authors thought that the material of the moon might have taken from an extremely thin surface layer of the earth, covering the area now being the Pacific. It is not necessary to discuss in details the impossibilities of this theory - fig.4 is sufficient to show it to be an absurdity. - Only concerning their ages the Pacific and the Atlantic may differ essentially. As one knows the existence of coral atolls in the Pacific, and their absence in the Atlantic, has been interpreted as a symptom

that the Atlantic might be considerably younger.

The purely hypsographical statements mentioned above have an analogon on Mars, where we also find two different levels of altitude. But this differentiation is less radical in the case of Mars: The altitudes of the two levels differ only by about 300 to 700 m, much less than the corresponding difference of about 5 km on the earth. The areas of the higher and of the deeper level have on the earth a ratio of about 2:3, on Mars 3:1.

The analogy of these features on two different but similar planets surely will become the object of much research as soon as expeditions to Mars will be possible. Today only the earth gives us opportunities to study this phenomenon of the two different levels in more details. At first its significance is emphazised by the additional fact, that the continental blocks are limited by steep slopes against the deeper level. Fig 5 gives examples of profils showing this clearly.

These statements - in their nature purely morphological, and concerning only the surface of the earth - now are completed by the results of research concerning the physical and chemical properties of the materials forming the continental blocks and the bottom of the ocean basins. Gravimetry, with its results showing isostasy as a rule with high significance - occasioned deviations from the rule having too their meaningful significance - allows

statements about the density of materials as well as about the thickness of layers of different densities. Seismographic measurements, giving knowledge about velocities of sound, as shown in earthquake waves, completes our physical knowledge of materials and their geometrical distribution. Petrographical analyse at last shows us the material of the continental blocks to be also chemically different from that of the deep sea.

Without going into details we remind the well known fact, that the material of the continental blocks containing a great part of granitic and similar rocks, is usually called SIAL , and that of the bottom of the ocean basins, prevailing of basaltic composition, SIMA. The densities of the two materials, though varying to a certain degree, are distinctly different. For instance measurement of veloci= ties of earthsquake waves as well as samples studied petro= graphically show that the material of the Atlantik Ridge surely is not sial, but sima, together with rock material from the depth, pulled out by volcanic activity. The strong earthquake activity along the Atlantic Ridge - in the vicinity of the rift - is not only proved for the present but also by age determinations of material found there; great parts of these basaltic rocks crystallised out of molten material only somewhat less than 10 million years ago, as shown by nuclear methods (potassium-argon method).

The ocean basins usually show only one or a few km of sediments, very little in comparision with the huge

masses of sediments on parts of the continental blocks. The continental blocks are layers of a thickness of about 35 or 40 km, and this thickness shows a high degree of constancy in the whole areas of continental blocks: Otherwise there could not be this high degree of isostasy, revealed empirically.

Already these facts - most of them well known since long years - lead to consequences not always recognised clearly enough. The differences not only concerning the level of altitudes, but also of chemical, petrographical composition make it impossible to believe that any mentionable transformation of continental areas into deep sea ever has taken place.

Take as an example the Gondwana land of the geologists. As far as I know there exists scarcely any disagreement among geologists as to the reality of Gondwanaland as a past broad bridge between Africa and South America. But this bridge cannot have existed in any other form than by geometrical vicinity of those masses of sialic material which also today form both the mentioned great continental blocks.

Therefore Wegener's well known theory of "Kontinental-Wanderungen" cannot be so false as it has been assumed to be during some time by most geologists. In the contrary some parts of the content of this theory must be acknowledged

as unavoidable truth.

The great difficulty of Wegener's theory in its original form was the lack of forces which would be able to cause the supposed movement of the great continental blocks on the sphere of the earth. This difficulty has been thoroughly discussed, from a physical point of view, by Jeffreys (24). His criticism indeed seems to show that the old form of this theory is physically untenable. But now apart from the considerations above palaeomagnetism gives strong support to the idea that the continental blocks performed great movements relative to each other, similar to what Wegener believed. Therefore it is a removal of difficulties otherwise scarcely to be overcome if now Dirac's hypothesis allows us to understand earth expansion as a sufficient cause to draw the different parts of the continental-block-layer from each other.

Another consequence of the fundamental facts about continental blocks and ocean basins is the impossibility of what has been called the growth of continents. Empirically one knows 11 so called kernels of continents ("Kontinentalschilde")- areas in which very old rocks (2 to 3 milliards of years) are lying at the surface, surrounded in a more or less concentric manner by other zones presenting younger rocks (for instance 800 to 1200 millions of years) at the surface - several such zones are following, so that the impression of a real "growth" of the continents during geological ages arose. But that must be a misinterpretation of the facts.

J.T. Wilson (20) and other authors are convicted that the growth of continents may be a real fact of the past. According to his theory the whole mass of sial originated in deeper layers of the crust of the earth, being brought to the surface by volcanic activity, leading to the formation of mountain ranges. This idea is improbable already because the material brought to the surface by strong and great volcanic eruptions is not of the sial type, but belongs to those other types of rocks forming deeper layers, which only by volcanic activity become occasionally visible for us; to assume granitic layers below the basaltic ones of the sima seems to be scarcely plausible. The atlantic ridge shows very clearly that volcanic activity in the deep sea does not produce sialic matter - we mentioned already this modern knowledge about the material of the Atlantic Ridge, which shows also that the latter cannot be interpreted (as has been done at the basis of Wegener's theory) as sialic material left back from the separation of Africa and South America. Another point of severe criticism against this theory of J.T. Wilson is, that Mars shows too the phenomenon of two different hypsographic levels, though volcanism (and mountain ranges) are absent there.

The existence of continents and ocean - or more precisely the existence of a sial layer, "swimming" in the more dense sima layer, possessing nearly constant thickness, but covering only 2/5 of the surface of the earth - is surely one of the great problems of natural sciences. Two theories trying to give an explanation have already been mentioned

above: The phantastic theory taking reference to the hypothetical separation of the moon from the earth, and secondly the theory of J.T. Wilson against which severe arguments of criticism must be put forward. A third theory, by Hills and Jeffreys (24) may only be mentioned as existing. It seems to me to be scarcely convincing, and its applieation to the corresponding features on Mars has never been discussed. Some authors today would prefer to believe that slow convection currents in the mantle of the earth might have caused that the material of the continents is united in a layer covering only 2/5 of the surface. But this is not a theory but only an idea or an hope; and an idea which scarcely can be brought into harmony with the constant thickness of the continental blocks.

The "theory of strong expansion of the earth " explains the phenomenon in the following manner. (J. Fisher, published 1952 in my book; L. Egyed 1957; B.C. Heezen (1960).) The sial layer at the time of its original formation - already with its present thickness - surrounded the whole earth, which therefore at that time must have had a radius which amounted to about $\sqrt{2/5}$ times the present one. The growth of this radius tore the sialic layer to several pieces, identical with present continental blocks. Fig 6 , according to Heezen, shows schematically the process of formation of the Atlantic, including the ridge.

This evidently explains the facts which have been misinterpreted as "growth of continents": The water of the sea must have covered great parts of the continental blocks when the area of the ocean basins was still considerably smaller than today.

Egyed from a careful study of the geological charts from Ternier and of those of Strahow gave a very convincing proof that this so called "growth of continents" persisted also during the last 500 million years, from the cambrium till now. His results are shown in fig 7 . We see here - according to his interpretation of the facts - that those parts of the continental blocks which were covered with water decreased from about 50 % or even 80 %, 500 million years ago, to practically zero today.

Additionally it may be mentioned that sediments with an age of about $2 \cdot 10^9$ years contain scarcely any sandstone. This seems to be understandable from the idea that at that time perhaps all those parts of the continents which are not higher than 2 km may have been still covered by water, so that there existed no sufficient areas to assort fragments of rocks in order to produce sand.

The interpretation of continental blocks and ocean basins according to the indicated idea of Fisher, Egyed, Heezen wins support from well known facts about the "Inselgirlanden". This phenomenon - seen for instance at the eastern coasts of Asia - has been cleared convincingly by seismic research. Such an arc of islands is lying in the cut of a certain cone with the surface of the earth - the cone having an axis of radial direction, and the point of the cone lying in depth's up to 500 km.

That segment of the cone which is indicated by the arc of islands is an area where numerous centres of earthquakes (Erdbebenherde) are situated: It is understood at the basis of

thorough geophysical research that these earthquakes are caused in such a manner that "Scherungs-Spannungen" in the sima cannot exceed a certain maximal value without giving rise to an earthquake. The tension along this segment of the cone area means that the matter inside this segment slowly is raised and put upward relative to the sima outside of the cone segment. This interpretation of the situation must be acknowledged as well founded and proven.

Most authors have seen in these facts a confirmation of contraction of the earth. But this conclusion ceases to be convincing if we try to give an explanation also for the further fact that the points (Spitzen) of these cones use to ly below the continental blocks, near their margins, so that the arc is concave to the continent and convex to the ocean. This law is not explained by the facts and ideas mentioned above; but it can be easily understood at the basis of expansion theory. Expansion of the earth with growing ocean basins, but constant area of continental blocks necessitates a certain flux of sima material from below the continental blocks to the ocean basins - where the surfase of the sima is situated only about 5 km deeper than that of the continents, against about 40 km there where the continents are swimming in the sima.

Only with this supplement the geophysical theory of island arcs becomes really convincing - and at the same time a new support for the expansion theory. But it must be mentioned that T.J. Wilson put forward the idea that all greater mountain ranges might be arcs of a similar character as

the island arcs. He gave no indication of any idea to explain such a phenomenon at the basis of physical laws; his purely morphological endeavour to show that geography of mountain ranges may be described in this manner does not seem to me to prove anything.

Also the study of palaeomagnetism seems to give support to the expansion theory of ocean basins. At first it may be mentioned that palaeomagnetism seems also to show the fact that the two magnetic poles of the earth in the past repeatedly performed a permutation - the magnetic field of the earth diminishing towards zero and then reestablishing itself with permutation of the poles. This is in harmony with the fact, that also the sun showed such permutation, and other magnetic stars too - perhaps stars with a constant magnetic dipole do not exist, and all endeavour to explain their existence as a consequence of known physical laws has been useless. The problem to explain variable magnetic stars has, as far as I know, not yet been discussed theoretically. Maybe this problem may lead to a decisive proof in favour of the theory of varying $\varkappa$; but nobody knows anything about this point.

Apart from this mentioned permutations of the magnetic poles of the earth palaeomagnetic measurements show continous variations, in the past, concerning the magnetisation of different parts of the continents. Endeavour to give an explanation of the empirical facts has lead to application of the often discussed idea of wandering of the poles. This idea, though many authors are inclined

to believe in it, is in reality not very plausible. Some authors are convinced that the poles during some geological times were situated at places in the tropical zone - it seems to be never realised that in such a case also the polar flattening of the earth must have made a corresponding "wandering", meaning a considerable deformation of the earth.

But immediately from the palaeomagnetic data it became visible that any hypothesis of pole wandering is not suitable to interpret the facts. For the attempt to reconstruct from these dates the way along which the north pole must have made its wandering gave totally different results if one used data from Europe, India, Australia, North America - the beginning of the reconstructed way must have been situated in the pacific, in South America, in Africa and again in the Pacific, but far away from the other point (won from the European data).

But Heezen announced that by a systematic reinterpretation of all available data came to the result that not pole wandering is necessary if one assumes expansion of the earth, strong enough to explain the ocean basins as indicated above, and separation of the continents in consequence of this expansion. Details of this important discovery of Heezen are not yet published. But perhaps they will settle the matter definitely.

As a last support to the theory discussed here it may be mentioned that erosion today is strong enough in order to have given in the course of 3.10^9 years a mass of sediments amounting to about the total mass of the sial. (20). This

would give a severe paradox if not according to the theory discussed above, in the past - with much smaller land areas than today - erosion must have been too considerably smaller.

H. Flohn [1] recently gave a very instructive discussion of a great part of the problems discussed in this chapter here. He did not yet take into consideration Dirac's hypothesis so that his results cannot be comparable to ours; but in many respect they are apt to give support and also supplements to our considerations here.

Intrusions and volcanism.

§ 5. H.J. Binge gave what I think must be regarded an important contribution to a better understanding of volcanism. Indications of his ideas are given with his kind permission in the second edition of my mentioned book (1) and in my two articles about the expansion of the earth (2). A full representation of Binge's idea will be published in the future. Here I include a sketch of the main points.

As already said above, Mars does not show any volcanism; and concerning our moon Baldwin's thorough discussion seems to me to have proven that the "craters" of the moon are not of volcanic origin, but results of the impact of numerous m meteoritic bodies.

The circular symmetry of these craters and the absence of lava production are facts easily to be understood at the

1) Naturwiss. Rundschau, 1959, page 375.

basis of the impact theory, but not explainable by hypotheses of eruption. Nearly in all cases the mass of the circular wall equals the deficit of mass in the interior. The distribution of craters is determined by chance - showing not, as volcanos at the earth do, lines of weaker resistence. The central cones which are to be found in many craters have been explained by Schardin as result of "implosions" following the explosions caused by the impact. Recently the occasional occurence of gaseous explosions on the moon has been proved. But these seem to be phenomena entirely unconnected with the processes which caused the formation of moon craters. Summarising these hints we seem to be justified to say that the moon as well as Mars does not show any phenomenon comparable with our volcanism.

We do not yet know whether Venus and Mercury have volcanoes; but what we know is certainly favourable to Binge's conviction that among celestial bodies of the type of the inner planets possession of volcanoes is dependent upon a minimum mass - this lying between those of Mars and the earth. Therefore Mercury should have no volcanoes; but Venus probably must have them - nearly to the same amount as the earth. According to Binge this situation means that gravitation is a determining factor in volcanism; and without using Dirac's hypothesis no correct interpretation of volcanism can be given.

Geology knows as a frequent occurence magmatic masses coming from layers below the crust and having risen into higher layers, in many cases (though not all) even to the sur= face of the earth. Volcanoes are only a special case of this more general process of ascending magma - in volcanoes the rising matter comes into the open, whereas in the more numerous cases of "intrusions" ("plutons"; "subvolcanoes") the ascending masses have been stopped below the surface. Though not all volcanic material is coming from great depths, a series of rocks as dunite, sepentinite, eclogite would be unknown to us without volcanism and intrusions: Coming from the depths, they would never been found in regions attainable to man without these processes bringing material from the depth to or near to the surface.

Volcanoes at Hawai and at Iceland, putting out basic lava, show that isostasy does not mean any hindrance against such processes. The kimberlit (with diamonds in it) made its rising process in an explosive manner, forming "Durchschlagsröhren". Intrusions not attaining the surface often have the result of raising great masses above them. For instance "bysmalithes", pieces of rocks from the depth rising in zylindrical or conical form, can raise by masses. In the Yellowstone-Park layers from Cambrium to Carbon have been raised by bysmalithes. Many further examples show the huge forces driving such rising magmatic matter. The russian geologist <u>Fersmann</u> wrote about a bysmalith at Cola: "The mechanism of intrusion can only be interpreted as <u>explosive eruption."</u> Also the laccolithes, showing spreading of of the magma through canals and grooves of

the rocks, can raise the layers above. Especially explosive seem to have been the processes forming "necks"; the layers above the rising material has been thrown away. But in some of these cases the eruptive material has not the character of tuff, but of massive rocks from the depth - giving the impression that not gaseous matter has been active in the explosion, but only an increasing volume of rocks caused the process.

This last remark leads us already to Binge's decisive conclusion: The primary process of rising magma is a phase transformation of rock material, leading from more dense material - showing stability at higher pressure - to more voluminous material, showing stability at lower pressures. Rittmann (10) who already recognised clearly the explosive character of volcanic or intrusional activity, emphazised "Entgasung" (separation of solid or fluid matter from gases contained in it) as a decisive factor in magmatic explosions. It is not necessary to deny the importance of this process emphazised by Rittmann; but the other possibility hinted at by Binge probably is of still greater importance. Binge (still unpublished) gave interesting details about petrographic processes leading to a decrease of density, as assumed in his theory of intrusions and volcanic activity. But I prefer to discuss here only his chief ideas. He comes to the conclusion that during geological evolution a monotonuous decrease of gravitational pressure must be acknowledged as primary cause of intrusions and volcanism.

Indeed these magmatic explosions show that the earth - in its crustal and subcrustal layers - is in a condition of instability; and there exist only two logical possibilities: 1) To assume that this instability is produced by processes during the formation of the earth, several 10^9 years ago; or 2) to ascribe this instability to a cause working without interruption since the formation of the earth, also today; and causing a slow change of the physical conditions determining the structure of the earth.

Any hypothesis in the direction 1) necessarily leads into scarcely tenable physical ideas. Also Rittmann's endeavour to present a theory of the origin of planets which could give a basis for his interpretation of volcanism, seems to be to me not convincing. Thinking about the possibility 2) there surely remains Dirac's hypothesis as the only idea hitherto discussed of this kind.

Systematical decrease of $\varkappa$ must lead to a transformation of rocks from high pressure phases to low pressure phases. In the case of an infinitely slow decrease of $\varkappa$ the rocks would be at every time in that phase which shows stability under the corresponding pressure. But the decrease of $\varkappa$, though very slow, is already fast enough in order to guarantee that the phase transformations could not go on fast enough to give continuous accomodation to the decreasing pressure. It is well known that phases of solid material can exist very long times as metastable phases under conditions of pressure and temperature which give to other phases the absolute

stability. Therefore it is a very natural conclusion that accomodation of the rock layers to the decreasing pressure is performed with delay, and then in the manner of explosive transformations from instability to stability. Afterwards, having seen the ability of Dirac's hypothesis to make volcanic activity and intrusions understandable, one may be inclined to say that without this hypothesis the facts of intrusion and volcanoes would be unexplainable indeed.

Binge sees a proof for an effective decrease of gravitational pressure still during the last 500 millions of years in the fact that in some districts geology can show that in the course of geological periods volcanism put upwards magmatic materials from increasing depth. That must not mean that the decrease of $\varkappa$ itself has still been considerable during these 500 million years: It can also be interpreted that that the expansion of the whole earth - delayed too by conditions of metastability - has been performed partially during these last 500 million years in supplementary accomodation to a decrease of $\varkappa$ performed already in former periods.

This last point will be discussed further in the beginning of the next chapter.

At the end of these considerations the great problem of the formation of mountain ranges may be mentioned - objet of so many conflicting theories. Surely a clear judgement

of this problem has been made more difficult than necessary by the circumstance that even the pure facts of the matter have become controversial - theoretical or hypothetical ideas having influenced also the judgment concerning the facts. Numerous authors have been inclined to speak about periods of mountain folding and periods of absence of folding ("Faltungszeiten" und "Ruhezeiten") as if they were talking about proven facts - instead of acknowledging all these ideas of world wide periodicity concerning formation of mountain ranges as purely hypothetical - and certainly false, as we know now. Naturally in any smaller geographical district there have been alternatingly periods of tectonic activity and inactivity. But all ideas of world wide alternating periods of this kind - or even periodicity in this respect - are entirely speculative ones, in clear contrast to the reality. This has been made probable already by the critical discussion given in the beautiful and instructive little book of Simon;[1] and it has been proven beyond any doubt since methods of nuclear physics made it possible to get systematic geological information concerning the last 3.10^9 years, instead of the 5.10^8 years of older geology. A very instructive analyse of the new situation concerning geological theories has been given by Wilson, Russel and McCann-Farquhar (16). Summarising we state here only that the process of formation

---

1) W. Simon, Zeitmarken der Erde. Braunschweig 1948.

of mountain ranges has been in activity during the last $3 \cdot 10^9$ years without interruption.

Surely this process - connected with eruptive processes, but probably in such a manner that both processes are to a certain degree independent from each other - shows severe complications, which to discuss in details is beyond the scope of this Report. But it may be emphazised that the theory of earth expansion contributes one totally new idea to this discussion: The continental blocks must have been deformed in a bending process, resulting from the increase of the radius of the earth. Perhaps this is exactly the governing factor of mountain folding.

§ 6. Palaeoclimatology. As already mentioned above, according to Dirac's hypothesis the radius of the orbit of the earth must have been smaller in the past. Therefore the earth received in the past more radiation from the sun.

Going into the same direction, but much more important is another consequence of Dirac's hypothesis: The production of radiation by the sun according to astrophysical theories is proportional to $\varkappa^{\frac{15}{2}}$, as shown in the well known astrophysical formula

$$\mathcal{L} \sim M^{\frac{11}{2}} R^{-\frac{1}{2}} \varkappa^{\frac{15}{2}} .$$

The radius R being too dependent upon $\varkappa$, what one calls the "solar constant" must have been during geological

ages proportional to about $\varkappa^9$. That means that the solar constant has been greater than today:

about 2% one million years ago;

about 50% during the carbon.

These consequences of Dirac's hypothesis lead us to an entirely changed picture of climatic development in the past. It has to be discussed carefully whether the known empirical facts are in accord with these strong consequences.

At first let us consider the palaeozoic aspects. Teller (6) who has been the first one to emphazise the importance of the $\mathcal{L}$ formula above in connection with Dirac's hypothesis, was inclined to interpret the mentioned consequence of this hypothesis as a proof that the hypothesis must be false: Organic life in palaeozoic ages would have been impossible under such conditions of solar radiation. But ter Haar (13) put forward the idea that such amounts of solar radiation might have caused the earth to be covered by closed strata of clouds. As one knows, Venus possesses such closed layers of clouds - some authors believe them to have a ~~th~~ thickness of the order of 100 km. Now the solar constant relative to Venus is about 4 times that relative to the earth; and conditions on Venus are therefore considerably more extreme, than on the palaeozoic earth - according to Dirac's hypothesis. But it seems at least plausible that an increase of the solar constant to 150% or 200% of its present value would

cause not a great increase of surface temperature, but in the first line an increase of cloud masses.

Now there are empirical facts favouring decidedly this idea. R. Potonié (14) gave a very suggestive discussion of climatic conditions in the carbon. These conditions are to be characterised as an astonishing homogenity and uniformity of a humid climate free from frost, with constant abundance of rain. (Coal-layers are extended from the Antartic continent til Spitzbergen). The humid air, frequent showers of rain, perhaps also hail, and the wide districts of swamps and water covered areas have been combined with extreme uniformity of the climate in its geographical distribution as well as in its temporal course. The trees have no annal rings ("Jahresringe"), a fact proving smallness of seasonal variation. Temperatures of 10 to 12° Celsius seem to have been maintained - without great variation - in great parts of the continents.

Already these facts obviously are in best accord with the idea that closed layers of clouds - radiated by the sun about 50% stronger than today - caused this uniform and extremely humid climate below them. But according to Potonié there are still more facts of high significance. Botanical details prove the carbon flora to have been a shadow flora. Ferns are still today shadow plants - unable, according to experimental studies, to be stimulated by periodic light effects. Potonié himself concluded that the structure

of the carbonic woods themselves must have caused deep, dark shadow in them - comparable to tropical virgin forest today. But he then encounters the difficulty that high trees of the carbonic woods gave only very little shadow. Summarising we may be allowed to say that the empirical facts about the climate of the carbon not only are in accord with the idea of closed cloud layers, but seem to prove the correctness of this idea.

But, as emphazised by Lotze (15), great stocks of salt are known also from periods before the carbon (cambrium to devon). Therefore drying of considerable areas of shallow sea must have been performed already during these periods. Taking regard to Dirac's hypothesis this fact perhaps must not necessarily be interpreted as result of great areas having no clouds during long times. Also districts with a relatively thinner layer of clouds may have been free during long times from rain, in consequence of a solar constant amounting to 150% or 200-250% of its present value.

A famous problem of palaeoclimatology is given by the palaeozoic glacial periods. Several of these are known, also in the devon[1] and in the perm. Conspicuous is the circumstance that the symptoms of this glaciation in the past are found chiefly along the equator and its vicinity , and everywhere in this zone: South America, Africa, South Asia, Australia.

---

1) O.H. Schindewolf, Akad. Wiss. u. Liter. Mainz 1951, S. 1

Many authors saw in these facts a proof of the correctness of hypotheses of pole wanderings: According to this interpretation the poles must have had in the past repeatedly places near the equator - and nearly each place along the equator must have been, at a certain time, the place of one of the poles.

Our endeavour to discuss the (unavoidable) consequences of Dirac's hypothesis leads to another interpretation. The closed layers of clouds which seem to be unavoidable for palaeozoic times must have had a maximum of thickness in the tropical zone. Therefore it might have been possible that not the poles but the equatorial zone has been in the past, several hundred million years ago, the pfeferred zone of glaciations: These huge masses of clouds, above what we now call the tropics, may have produced in some periods masses of snow not being able to be molten fast enough in order to avoid accumulation. Though seeming paradox at first sight, this idea certainly is not more paradox than for instance the well known fact that the stratosphere today has lower temperature at the equator than at the poles; and it may be mentioned also that some tropical high mountain zones today are covered by snow not in their winter, but in their summer season - for then the amount of falling snow is considerably greater (C. Troll).

This perhaps quite "revolutionary" interpretation of palaeozoic glaciations can perhaps give also a solution of another vexing problem. It is well known that there

these exist in the oceans "guyots", mountains with horizontal abraision areas ("Tafelberge"), showing that during long periods in the past the surface of the sea has been at the altitude of these abraisions. These are lying partly about 1500 m below the surface of the present oceans.

It is well known at the other hand that during the last glaciations, in the last 600 000 years, that the accumulation of huge masses of ice reduced the surface of the oceans to altitudes lower than now, to the order of 100 m. It seems to be scarcely possible to find any other explanation of the guyots than to conclude that there existed periods in the past during which huge glaciations subtracted water from the oceans in amounts even much greater than during the last, diluvial glaciations.

This conclusion, if correct, evidently favours the idea of equatorial instead of circumpolar glaciation - geometry causes equatorial glaciation, if real, to accumulate still much more ice than polar glaciation.

In an highly interesting manner Dicke (5) discussed the consequences of Dirac's hypothesis and the astrophysical formula above with special regard to the oldest times in the history of the earth. Backwards to about 2.10^9 years before today the closed layers of clouds gave a strong diminution of the surface temperature of the earth. But this temperature, 10^9 to 2.10^9 years ago, must have amounted

to 50 to 100° Celsius. In still older times the whole water, also of the present oceans, must have been in the atmosphere, but not in clouds but in a purely gaseous state without formation of drops. Though under these conditions the atmo= sphere must have been totally opaque, according to Dicke the surface temperature of the earth must have amounted to 1000° (or more) Celsius, probably not yet allowing formation of solid rocks. Also in this direction we see the revolutionary consequences of Dirac's hypothesis - but in harmony with those data which seem to make it probable that the oldest rocks are already considerably younger than the earth itself.

Our simple discussion above about palaeoclimatological consequences of Dirac's hypothesis remains far from any attempt to develop systematically and in a quantitative manner what must be deduced from this hypothesis. Any attempt of this kind would have take regard also to the following points:

a) The solar radiation in the past must have been not only more intense but also have shown a different spectral distribution. The surface of the sun, according to its much greater luminosity, must have had a higher temperature. But also the eruptive activity on the sun - influencing not only our aurorae, but being also meteorologically meanigful, must have shown differences relative to today.

b) With greater gravity also the air pressure at the surface of the earth - and the coefficient of its decrease with growing altitudes - must have been greater than now.

c) With a smaller radius of the earth the duration of the day must have been smaller - a point not without importance concerning finer details of meteorology. -

The author is thankful for interesting discussions, partly extended ones, partly short ones, with the colleages Becken, Binge, Dicke, Flohn, Lotze, Schwarzschild, Vieweg.

§ 7. The diluvial glaciations. More than 50 different published theories attempted to explain the diluvial glaciations; and none of them has been acknowledged generally as convincing.

Therefore it is surely interesting to see what consequences concerning this great and famous problem arise if we assume Dirac's hypothesis to be correct. Considering this point we may emphazise once more that the primary aim of all our discussions here is to clear the question, whether from Dirac's hypothesis any conclusion is to be drawn which would be a contradiction against proved facts. Till now we only saw that these consequences partly are quite revolutionary in their diversity from orthodox meanings, but that known empirical facts seem to be in accord with them. Putting the same question now about palaeoclimatology of the diluvium, we may take in mind that the many different details of the problem can interest us here

only so far as they may have any relation to the Dirac hypothesis. We are therefore not interested in possibilities to explain this or that by making assumptions of any kind - and we do not intend to show special ideas to be better than other ones fitted to give explanations of some established facts. All we intend to do is to discuss whether those consequences which arise unavoidably from Dirac's hypothesis are in accord with known facts or not.

That the diluvial problems have their own relations to Dirac's hypothesis is made clear already by the foregoing paragraph. For we saw there that the hypothesis of diminishing gravitation leads us to totally new ideas about the paleozoic glaciations - and this gives considerably changed aspects also concerning diluvial glaciations which hitherto have generally be thought of as a repetition of similar events as those which occured in the former glaciations. But now we see that the Dirac hypothesis necessitates the acknowledgement that the paleaozoic and the diluvial glaciations have been totally different phenomena - the older ones having been extended in the vicinities of the equator, the other ones in the vicinity of the poles. If this is correct, there is no need to look for such explanations of the diluvial glaciations which would be applicable also to the older ones: in the contrary the problem now takes the form: Why could in the present (the alluvium being probably only one of the "Zwischeneiszeiten") this phenomenon arise which has never in a similar manner taken place in all geologic past?

The basis of our discussion of the diluvium may be the supposition that the well known Milankowitsch curve really gives the correct temporal intervalls of the diluvial glaciations. The author of this report cannot say anything about the validity of this supposition. Several leading specialists of diluvial geology and climatology are (or have been) convinced of the correctness of the Milankowitsch theory; but there are others who are inclined to have doubts.

To take the curve - in spite of these doubts - as basis of our discussion seams to be justified because only then, if the Milankowitsch curve is really meaningful, a discussion of the diluvium in connection with Dirac's hypothesis is necessary. If the time table of the diluvium is not determined by this curve, then there exist many different possibilities to make attempts to explain the diluvial glaciations, and the Dirac hypothesis can be left aside in all discussions about this topic. For instance there is the possibility that clouds of interstellar matter may have reduced during some time the value of the solar constant. That may be a possibility - as one knows, some authors believe this to be the solution of the problem. But such a cloud surely could not have been exactly homogenous in its density - our sun, crossing such a cloud, would have been during this travel some time in more and some time in less dense districts. The alternation of glaciation and of melting phases during the diluvium must then have been determined by the density changes of the cloud along the path of the sun. And naturally any of the other more than 50 theories of diluvial glaciation may be correct, without any connection with Dirac's hypothesis. But if indeed the Milankowitsch curve gives the correct time table of the diluvium, then there exists a clear problem,

arising from a well defined but till now unexplained phenomenon. This we intend to discuss here.

To the convenience of the reader we at first summarize shortly what has to be said about the Milankowitsch curve and its (controversial) significance. The excentricity of the orbit of the earth varies between minimal and maximal values with a period of 91800 years. The perihelion, performing a slow wandering around the sun, is wandering therefore through the seasons, its period being 20 700 years. The angle between the axis of the earth and the direction orthogonal to the plain of its orbit varies between 22° and $24{,}5^{\circ}$ with a period of 40 400 years. These periods have been studied by astronomers partly already long ago, and Köppen and Wegener attempted to use their results as a tool of palaeoclimatology. Interesting is in the first line the temporal variation of solar radiation at the earth during the summer-glaciation being surely determined not so much by circumstances during the winter as by the chances of melting during the summer. Milankowitsch gave a more precise theoretical evaluation of the corresponding "radiation curve" with regard to the last 650 000 years; later even for 10^6 years. The minima of solar radiation in the summer seem to be in good accord with the geological statements about (each) two Würm-, Riss-, Mindel-, Günz-glaciation periods. - It is to be taken in mind that according to this theory glaciation periods at the northern and at the southern half sphere do not coincide, a consequence not yet already tested empirically as thoroughly as one might demand. But direct temporal comparision of diluvial events in the north and in the south is not easily to be accomplished - only the reduction of the surface altitude of the oceans by accumulation of ice masses mentioned already above being an effect of world wide significance.

Now if this interpretation of the time table of the diluvium is correct, then the situation is thus: We understand the temporal intervalls of the diluvial glaciations; but we do not understand why these glaciations took place during the last 650 000 years, though they did not take place - in corresponding intervalls - also in former geological periods. (Climatological conditions not too far from those of today have been proved to exist already in the time of belemnites, by isotopic analyse of tails of belemnites).

We must therefore conclude that the <u>primary cause</u> of the diluvial phenomena has been a reduction of the solar constant performed since about 10^6 years, and this reduction was strong enough to bring the <u>average</u> temperatures during these 10^6 years near to the <u>verge</u> of glaciations, in a manner which previously did not occur. It is probable that this reduction has been performed in a slow, continous process - the cold periods in the intervall from 650 000 to 10^6 years before today seem to have been less strong than the later ones; and glaciation periods concerning only the Alps seem to have occured also <u>before</u> 10^6 years.

But the chief and most interesting and astonishing feature is the monotonuous slowness of this reduction which evidently took place without any mentionable irregular variations - so that the change between glaciation and warm melting phases could be regulated by the weak influence of small variations in the orbital movement of the earth. It seemed till now extremely difficult to produce any idea to explain <u>such</u> a reduction - that <u>Nölke's</u> idea of cosmic clouds

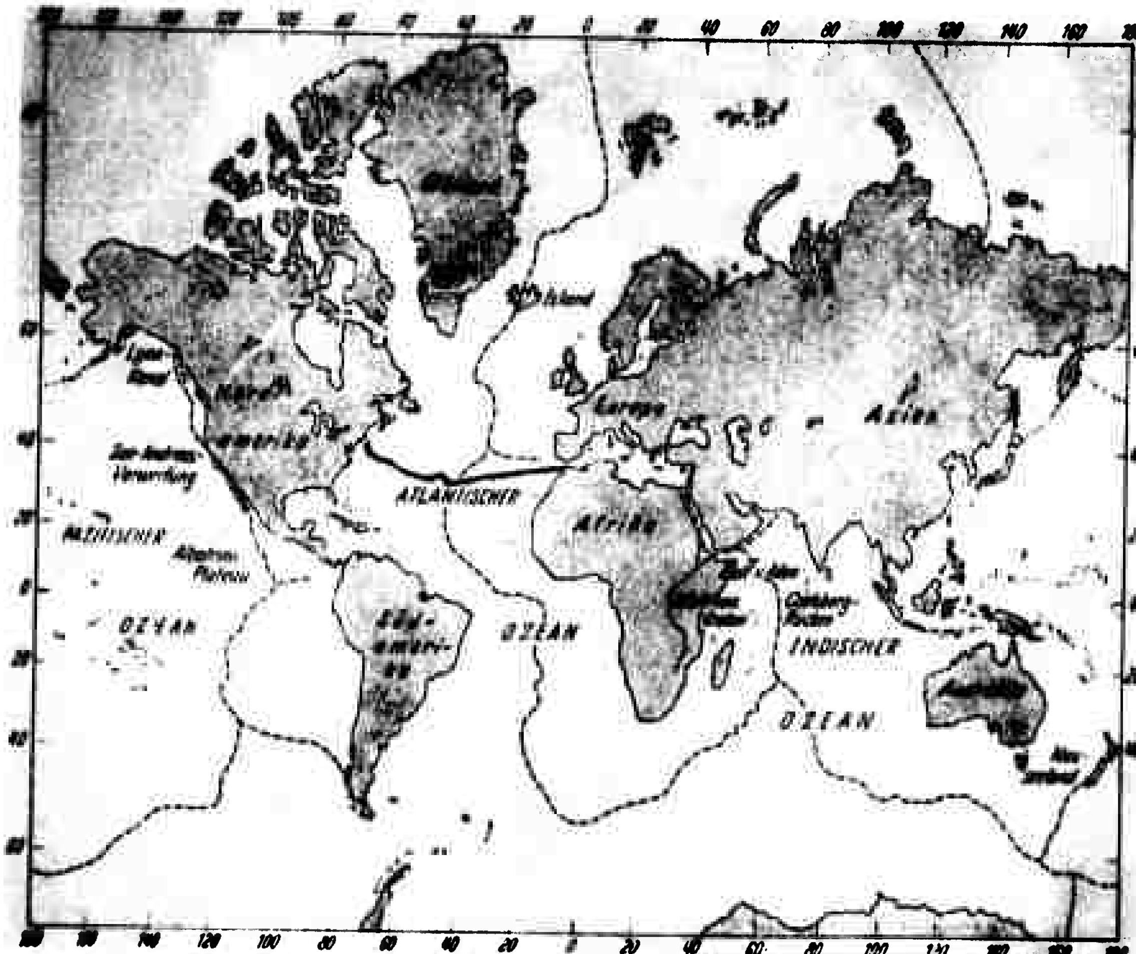


Fig. 1. The system of rifts (Ewing, Heezen, Tharpe)

Fig.2. Mjösen rifts (Stille)

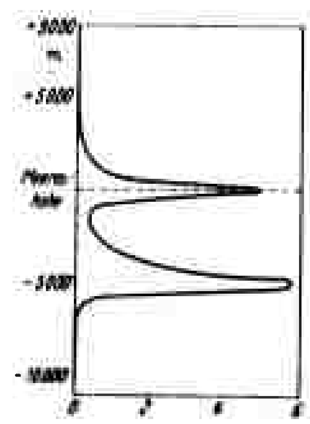


Fig. 3. Frequency of altitudes (Wegener, Bucher)

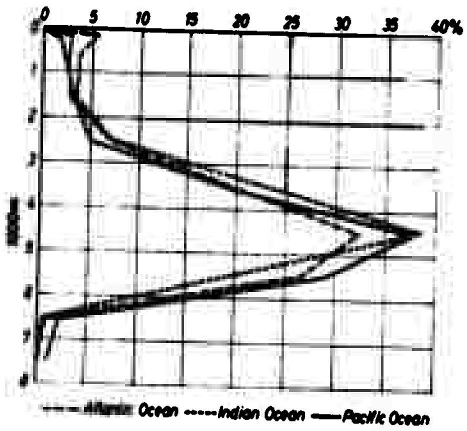


Fig.4. Frequency of altitudes in the different oceans (Egyed)

Fig. 5. Profiles of the continental slope (Heezen)

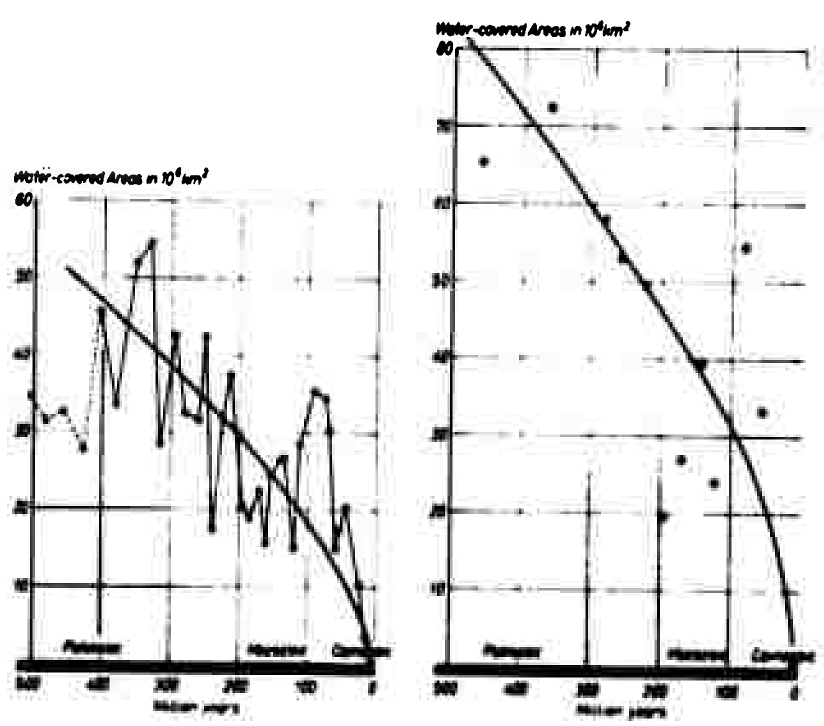


Fig. 7. Water covering of continents (Egyed)

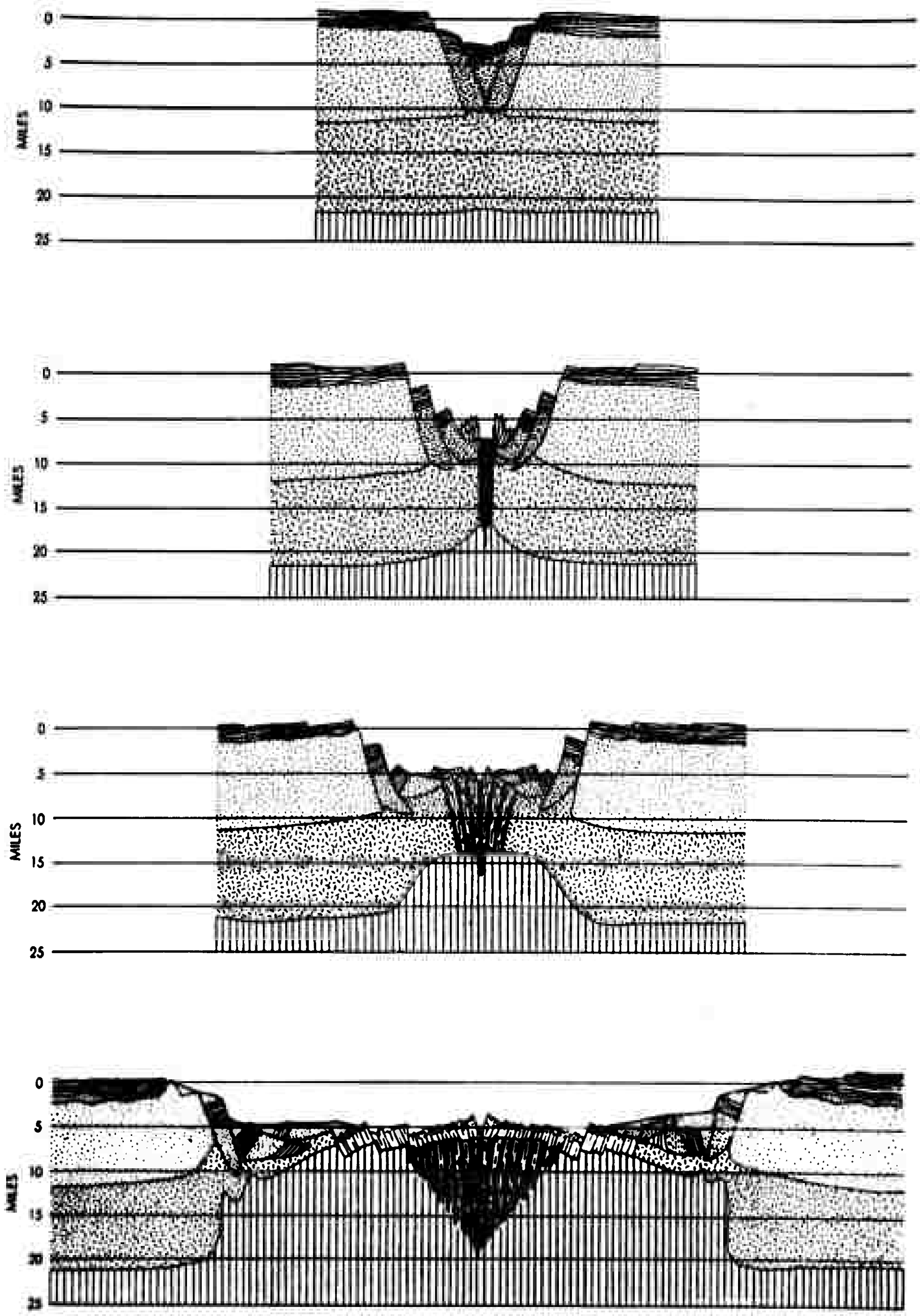


Fig. 6. Evolution of ocean bottom (Heezen)

is excluded if the Milankowitsch time table is correct, has been emphazised already above.

Therefore Penck has been inclined to criticise and abandone the theory of Milankowitsch: "So schön die Milankowitsch-Theorie auch erscheinen mag, richtig kann sie doch nicht sein; denn wenn das Eiszeitalter durch die Variationen der Erdbahnelemente verursacht wurde, so sollte es wohl auch im Mesozoikum und Tertiär Vereisungen gegeben haben. Die Bahnelemente waren immer gleichermassen veränderlich, aber von früheren Verei= sungen gibt es keine Spur". (Repeated here after Bacsák György). Other authors, defending the theory of Milankowitsch, have emphasized too that only the change of cold and warm inter= valls during the diluvium can be explained in this manner, but not also the fact that the diluvium in contrast to the tertiaer and the mesozoicum showed glaciation during its colder phases.

A highly interesting different theory has been put forward by Bacsák György. According to his opinion the finer details of the mechanical conditions in the system of planets are just now of a certain irregular or singular character, and the perturbations of the orbital movement of the earth, caused by the other planets, therefore could and must be such ones during the last million years, that similar perturbations did not exist since the palaeozoicum. Therefore he is convinced that Penck's criticism has been unfounded, and that no addition of another still unknown factor to the theory of Milankowitsch must be made in order to get a theory explaining to 100% all glaciations of the past only at the basis of celestial mechanics, without additional hypotheses.

This surely is a highly interesting matter. The author of this report feels himself unable to judge about this interesting theory of Bacsák György[1] which perhaps indeed gives the final solution of the whole problem. I can only emphasize the importance of the possibilities discussed by Bacsák György: It seems to me to be strongly demanded that his theory is studied by specialists, and that extensive new theoretical research in celestial mechanics is made with the aid of modern electronic computing devices. It would be useful already to make new computations extending the Milankowitsch curve over a period of 10^7 instead of 10^6 years.

If this new theory of Bacsák György is correct, then there does not exist any connection between diluvial phenomena and Dirac's hypothesis, and the whole matter of palaeoclimatology of the diluvium better had be left out of my discussion of the Dirac hypothesis. But in order to cover all possibilities we may consider here also the question what consequences would arise if future research would show the real situation to correspond more to Penck's idea than to Bacsák György's.

In this case it is clear that at least qualitatively Dirac's hypothesis is apt to explain exactly such a reduction of the solar constant as we saw to be required in order to fill out what is still lacking in the Milankowitsch theory. But also quantitatively one can see from the Milankowitsch curve that our statement above - about the order of magnitude of the reduction of the solar constant, to be expected theoretically - is in accord with the empirical facts. It seems

---

1) Bulletin of the Hungarian Geolog.Soc. 85, 70 (1955).

therefore to be possible that indeed the slow monotonuous reduction of the solar constant, caused by the decrease of $\varkappa$, is the primary cause of the fact that since about one million years the earth is at the verge of polar glaciation, in contrast to all former times.

CHAPTER II. ASTROPHYSICAL CONSEQUENCES OF DIRAC'S HYPOTHESIS

§ 1. The origin of double stars and of planetes. Famous theoretical investigations of Jeans [1] and other authors as Mc Laurin, Jacobi, Poincaré, Liapounoff, Darwin, Moulton, Cartan, Lyttleton [2] about the mechanism of division of rotating stars gave very significant results. Stars consisting of "fluid" matter can divide itself into two masses of equal order of magnitude, rotating about another in narrow orbits. But a "gaseous" star with a strong concentration of mass in its centre must perform, in the case of strong rotation, quite another development: It comes from an ellipsoidal shape to a lenticular one, wins a sharp edge, and begins to loose matter continually from this edge.

Though the pioneer work of Jeans, represented in the mentioned two great books, contains much details which today are not still up to date - owing to the rapid progress of modern stronomy - the mentioned chief results certainly seem to be of invariable value. The interesting book of O. Struve about "Stellar Evolution" [3] among its rich empirical material contains also ample proof that boths processes studied theoretically by Jeans really occur in the galaxis. Especially the division of a star into two stars is apparently

---

1) T. H. Jeans, Problems of Cosmology and Stellar Dynamics. Cambridge 1919. Astronomy and Cosmology, Cambridge 1928.

2) R.A. Lyttleton, The Stability of Rotating Liquid Masses. Cambridge 1953

3) Princeton 1950.

quite a normal process; there are examples of double stars showing the first stages after the division. It seems therefore quite sure that Jeans correctly interpreted at least those double stars which possess periods of the order of one day, indeed as originated by division in the manner analysed by Jeans. At the other hand there are examples also of the other one of both these division processes - with ejected matter collecting itself in an equatorial ring around the rotating star.

But the double stars with periods of the order of one day form only one of two extreme cases: A considerable part of all stars are double stars, showing all values of periods from about one day up to 10^3 years, in rare cases even more than 10^4 years. And the great majority of these pairs of stars corresponds to the condition that the mass relation of the two partners has the order 1.

Russel studied theoretically multiple stars of higher order. One knows examples of systems containing two components (wide apart), each one of these having two secondary components (narrow in comparision to the distance of the primary components), and at least one of these being a narrow double star indead of a simple one. Russel in a very suggestive manner established his conviction that these multiple systems are to be interpreted as results of repeated divisions of rotating stars. He put forward also the thesis that the whole phenomenon of double stars must be thought of as a uniform one, so that practically all double stars, up to periods of 10^3 or 10^4 years, have to be originated from the process of division.

In order to make this possible an effect is needed which could increase the periods of double stars - in extreme cases so that periods which at the beginning were smaller than one day could grow to more than 10^3 years.

Jeans himself heroically endeavoured to find such an effect. Tidal friction can cause an increase of the orbital radius of a double star. But it is to be conjectured that the order of magnitude of the period can scarcely be altered; and Jeans thoroughly proved and precisized this conjecture. Better working are perturbations of double stars by other stars; but Jeans by thorough investigation came to the result that these pertur= bations can cause an increase of periods only up to about 55 days - still greater periods are not increasing, but decrea= sing under the influence of these perturbations. Jeans discussed also the possibility of a diminution of stellar masses [by conversion into radiation] - showing in this manner how earnestly he endeavoured to find a possible cause of increasing periods of double stars. Our present know= ledge shows this idea to be untenable; and Jeans himself abandoned it.

He therefore saw himself forced to abandon also the thesis that the phenomenon of double stars would be an uniform phenomenon. Astronomers divide the double stars into two classes: The spectroscopic ones and the visual ones. Obviously this is a discrimination based solely at technical differences of observation and research. But Jeans - abandoning all hope to understand the double stars as an uniform phenomenon - put for= ward the conclusion, that this discrimination according to methods of observation by chance would coincide with a difference in nature, and that only the spectroscopic double stars - their periods indeed going up to about 55 days - might have originated by division. The visual ones might never have been single stars.

He detected also a special argument against the concept of Russel: The multiple stars of Russel in some cases have components with densities too low in order to fulfil the theoretical requirements: Theoretically they would have to perform their rotational division not by division in two partners, but by equatorial loss of matter.

This is the status of the problem of double stars - one of the most prominent problems in stellar astronomy. And it is obvious that Dirac's hypothesis gives us exactly what is needed in order to fill that gap of known laws of Nature which hampered Jeans in his endeavour to understand this fundamental feature of the realm of stars. That in a gravitational two body problem the orbital radius has to increase proportional to $\varkappa^{-1}$, has already be emphasized above. But also the special difficulty of too low densities of components of some double or multiple stars vanishes if Dirac's hypothesis is taken as a basis of the discussion: Greater values of $\varkappa$ in some cases will be apt to transform stars of low density into stars of much greater density.

At last it may be said that also the theory of the origin of planets, in the form put forward by V. Weizsäcker, seems to become tenable only at the basis of Dirac's hypothesis. According to ter Haar[1)] there arises the following difficulty: Starting from v. Weizsäckers ideas one can compute that the lens of gaseous matter surrounding the sun, in which v. Weizsäcker

1) D. ter Haar, Ap. J. 111, 179 (1950).
Nature 158, 874 (1946).

sees the material from which the planets have been built up, would be scattered into the surrounding vacuum already after 10^3 years. But 10^8 years would be needed in order to allow formation of planets out of this matter. Therefore the Weizsäcker process would have been impossible with the present value of $\varkappa$.

It may be remarked that the results of Jeans suggest a certain addition to v. Weizsäckers theory - increasing its convincing power. The lens of gaseous material around the sun postulated by v. Weizsäcker may have been more a ring than a lens - a ring originated by a rapidly rotating primeval sun performing the equatorial ejection of matter analysed theoretically by Jeans.

§ 2. Dirac's second hypothesis. For convenience of the reader the original argument of Dirac concerning his hypothesis may be recapitulated here, with some additional remarks.

The gravitational force in the action between a proton and an electron is about 10^{-40} times the Coulomb force between these two particles. Dirac felt it to be extremely difficult to believe in the existence of laws of Nature explaining a dimensionless constant of the orderer of 10^{40}. But also the age of the universe, if expressed in l/c with $l = 10^{-13}$cm ("elementary length") is of the order of magnitude 10^{40} . Therefore Dirac proposed the idea that $\varkappa$ might be a function of the age of the universe, being inversely proportional to this age.

Concerning this age some uncertainty arose in the last years since, as one knows, a few objects have been detected to which ages of 13 or even 24 billion years have been attributed. But the age determinations of these objects make use of the astrophysical luminosity formula discussed already above, $\mathcal{L}$ being strongly decreasing with decreasing $\varkappa$. Therefore, introducing the Dirac hypothesis, these huge values of ages are reduced strongly. Dicke[1] has shown that these and other discrepancies allow adjustment to each other if Dirac's hypothesis of a variable $\varkappa$ is made the basis of discussion. An age of the universe not exceeding considerably what has been thought to be correct several years ago can also today be judged as probable.

But there remains a difficulty concerning the expansion of the earth. If this expansion has been strong enough to make explainable also the ocean basins, then $\varkappa$ must have been rather great during the formation of the earth; and in this case the consumption of H by the sun, according to the luminosity formula, must have been very great - greater than in accord with the present chemical composition of the sun.

Concerning the expansion of the earth we learned above that Binge concluded that the expansion rate during the last few 100 million years might have been greater than it would have been in the case of equilibrium expansion, with a radius corresponding at every time to the value of $\varkappa$ at that time. It seems, that expansion shows a certain delay or relaxation compared with the temporal development of $\varkappa$.

---

1) R. H. Dicke, Implications for Cosmology of Stellar and Galactic Evolution Rates. Princeton 1961

This delay, as already said above, could be understandable as a result of phase transitions performed in the course of expansion; and it may be emphasized that a series of well known discontinuity spheres in the interior of the earth, in depths[1] of 413, 984, 2898, 4982, 5121, 6371 km, suggest the interpretation that they partly are not discontinuity spheres of chemical composition, but only of phase stability.

W.H. Ramsey [2] especially believed that the famous discontinuity at 2898 km might be a phase discontinuity instead of a chemical discontinuity; but concerning this case discussion with E. Teller convinced me that Ramsey's idea cannot be correct.

Concerning the other cases, interpretation of density discontinuities as phase discontinuities obviously has the consequence to increase the amount of expansion which we can await as consequence of decreasing $\varkappa$. But it seems that the difficulty indicated above can scarcely be overcome as long as we are inclined to think that the temporal development of $\varkappa$ in the vicinity of the sun has been the same as that in the whole universe. But acknowledging $\varkappa$ to be a fuction of time and space we can also believe that $\varkappa$ in the stages of formation of the earth has been different in the vicinity of the sun from its average value in the universe; and obviously the difficulty mentioned above would vanish if we could make it believable that $\varkappa$ has been quite great during the formation of the earth

---

1) T.A. Jacobs, Handb. d. Physik XLVII, 364 (1956)

2) Monthly Not. Roy. Astr. Soc. 108, 406 (1948)

but afterwards rapidly accomodated its value to the spatial average value in the universe.

But again this seems difficult without another radical step of thought. How can $\varkappa$ remain in any limited district considerably different from the average value, during times certainly greater than the radius of this district divided by the velocity of light? The answer seems to be: Only geometrical conditions - strong deviations from a flat, practically quasi - Euclidian Minkowski-world - can allow such an occurence. Therefore we become now strongly interested in a branch of relativistic research emphasized especially by Wheeler 1): Topology problems in General Relativity. Wheeler himself and his collaborators studied this field because it might be interesting with respect to the theory of elementary particles. But here we are interested in this topic because it may be meaningful with respect to stars and to the theory of formation of stars.

There is still another way leading to the same conclusion that topology of Riemannian spaces or topology of curved Minkowski worlds could give the clue to the problem of origin of stars.

The arguments of Dirac, as indicated above, can also be used in application to the density μ of matter in the universe. This average density, compared to the density of

---

1) Compare the lectures of D. Brill about General Relativity, delivered at the State University of Iowa, 1961.

nuclei[1], too is of the order 10^{-40}.

Therefore μ as well as $\varkappa$ might be inversely proportional to the age of the universe; and in combination with the Hubble effect this means that there must occur a process of generation of matter. Indeed Dirac 1937 believed that such a process must be assumed; but he later abandoned this idea. The well known "steady state theory" at the other hand assumes such generation, but in another form than according to Dirac's "second hypothesis" in its original form.

Now field equations belonging to a variational principle as the fundamental principle (5) of the "generalised gravitation theory" can never give production of matter - this point, not yet judged correctly by the author in his book, has been cleared later by discussion and criticism of Pauli, Fierz, Lichnerowicz. The present status of affairs has been indicated by the author in an already mentioned note [2] two years ago. But topology of curved Minkowski worlds offers possibilities to harmonise also Dirac's second hypothesis with the generalised theory of gravitation, and makes the way free to a new theory of formation of stars.

---

1) This means, that as in the case of $\varkappa$ we consider the physical constants c,h,l, as the "natural units" of physics.

2) Z. Physik 157, 112 (1959).

Nearly all those numerous astronomers and physicists who speculated about processes of formation of stars and galaxies tried ~~the~~ common, ~~the~~ conventional idea of condensation of interstellar gas. Oort, Spitzer and some other authors indeed came to results which seem to be suited to make it probable that a certain type of condensation leading to the formation of stars really occurs. But this must not mean that also the majority of stars originated in this manner.

Ambarzumian who studied the problem, not starting from conventional ideas, but starting from a penetrating and convincing analysis of empirical facts came to other conclusions. According to him the chief process to be revealed here is not the formation of stars from gas but the common formation of both kinds of stellar matter - gas and stars - from "proto stars" of "prestellar matter". This thesis, established at first by the extensive study of "associations" of stars, has been extended by Ambarzumian to galaxies and systems of galaxies. In the case of associations we have groups of stars (in some cases relatively young ones) showing by their motion to have originated from one point. In the case of galaxies there are many examples showing that the centre of a galaxy may be a field where the process of formation of gas and stars from prestellar matter is still in activity. This can lead to a division of the galaxy, or to other effects discussed by Ambarzumian.[1)]

According to these ideas the physical meaning of prestellar matter must be evaluated from the further analysis of the empirical facts, augmented now rapidly with the progress of

---

1) Ambarzumian, Solvay Congr. 1958 Brussels.

astrophysical research. Certainly this prestellar matter must have the following properties:

1) High density.
2) Invisibility in consequence of absence of luminosity.
3) Tendency to expansive or explosive outbursts leading to the formation of normal matter like gas and stars.

It seems to be quite difficult to detect any possibility to understand these properties by application of current physical theories. There must be conditions preventing the prestellar matter - before its transformation to normal matter - from emission of radiation; and there must be conditions allowing to maintain the high density of this prestellar matter before the beginning of the expansion leading to (or equivalent with) the transformation. But there must also remain the possibility to perform these expansive or explosive processes.

The most convincing answer seems to me to be that prestellar matter - as long as it remains prestellar - is prevented from emission and from explosion by geometrical conditions; that space has "pockets" containing the prestellar matter. Obviously in the frame of Riemannian geometry qualitatively there does not arise any difficulty if we conceive such pockets, separating the prestellar matter contained in them, but varying geometrically in the course of time so that at last the prestellar matter is poured into the normal, nearly euclidian space.

But a quantitative theory of such developments cannot be given in the frame of Einsteins theory of gravitation. For in this theory there exists the well known theorem of Birkhoff, according

to which any spherically symmetric vacuum field is necessarily a static one. This is a very important point: The conversation law of matter $T^{kl}{}_{\| l} = 0$ does not contain any difficulty against the indicated "pocket theory of prestellar matter". Naturally conservation of matter means now conservation of the sum of normal and prestellar matter. But Birkhoffs theorem really prevents us to develop such ideas inside the frame of Einstein's theory; for in a certain approximation each point in the astronomical space is surrounded by a spherically vacuum field.[1)]

But these difficulties seem to vanish if the Einstein theory of gravitation is replaced by the generalised theory of gravitation. Surely the questions arising in this connection will demand still much hard research work. But it seems that they are promising.

---

1) This difficulty can, so it seems to me, not be removed alone at the basis of M. Kruskals highly interesting discussion of the Schwarzschild solution.

Literature

(apart from those books and articles mentioned already above in footnotes of the text).

(1) JORDAN, P.: Schwerkraft und Weltall. Braunschweig: Vieweg & Sohn 1952.; zweite Auflage 1955.-

(2) JORDAN, P.: Akad. Mainz, Naturw. Kl., 1959, S. 771.-

(3) EGYED, L. : Geol. Rdsch. 46, 1o8 (1957). Vgl. ferner: Geofis. pura e appl. 45,115 (1960). Acta Zool. Acad.Hung. 1957,S. 201.Ann.Univ.Sci.Budapest, Sect.Geol.1957, S.37;1959, S.89

EGYED,L., u.
STEGENA,L.: Z.Geophys. 24,260 (1959).-

(4) HEEZEN,B.C.: Sci. Amer. 203,98 (1960).-

(5) DICKE R.H.: Rev. Mod. Physics 29, 355 (1957).-

(6) TELLER, E.: Physic.Rev. 73, 801 (1948).-

(7) WILSON,J.TUZO: Nature (London) 185,880 (1960).-

(8) EWING,M., u.B.C. HEEZEN: Amer. Geoph.Union Monog.1,75 (1956).-
EWING,J., u.M.EWING: Bull.Geol.Soc.Amer. 70,291 (1959).-
HEEZEN, B.C.,M.THARP u.M.EWING: Geol.Soc.Amer.,Special Paper 65 (1959).-

(9) STILLE, H.: Festschr. Akad.Göttingen 1951,I.S.71.-

(10) RITTMANN,A.: Vulkane uns ihre Tätigkeit, 2. Aufl. Stuttgart: Ferdinand Enke, 1960.-

(11) BALDWIN,R.B. The Face of the Moon. Chicago: The Univ. of Cicago Press 1944.-

(12) JORDAN, P.: Z. Physik 157,112 (1959)

(13) HAAR,D. TER: Rev. Mod.Physics 21,119 (1950).-

(14) POTONIÊ,R.: Naturwissenschaften 40,119 (1953).-

(15) LOTZE,F.: Steinsalz und Kalisalze. I. Teil. Berlin: W.Born=trager 1957.-

(16) WILSON,T.J.,R.D.RUSSEL u. R.McCANN FARQUHAR: Handbuch der Physik, Bd. 47, Geophysik I,S.288.Berlin-Göttingen-Heidelberg: Springer 1956.-

(17) SCHWARZBACH,M.: Das Klima der Vorzeit. Stuttgart: Ferdinand Enke 1950.-

(18) DURHAM, J.W.: Palaeoclimates. In: Phys. and Chem. of the Earth, Bd.3, S.1. London-New York: Pergamon Press 1959.-

(19) DEFANT, A.: Handbuch der Experimentalphysik (hrsg.v. W.WIEN u. F. HARMS), Bd. XX, Geophysik II, S.575. Leipzig: Akad. Verlagsges. 1931.-

(20) WILSON, J.TUZO: The development and Structure of the Earth. In: G.P. KUIPER, The Earth as a Planet,S. 138, Chicago: The Univ. of Chicago Press (1958).-

(21) JACOBS, J.A.: Handbuch der Physik, Bd. 47, S. 364. Berlin-Göttingen-Heidelberg: Springer 1956.-

(22) DAVID,E.: Z.Geophys. 25,239 (1959).-

(23) DIRAC, P.A.M.: Nature (London) 139, 323 (1937).-

(24) JEFFREYS,H.: The Earth, 4. Aufl. Cambridge: University Press 1959.

PART II. PROGRESS IN THE MATHEMATICAL DEVELOPMENT OF EINSTEIN'S THEORY OF GRAVITATION.

CHAPTER III. THE EXPANSION-FREE RADIATION FIELDS

§ 1 Introduction.

In recent days the problem of gravitational radiation has been attacked by a great number of physicists, in the hope to get a better insight into both the mathematical structure and physical interpretation of EINSTEIN'S theory of gravitation. Even though the hope of really measuring a gravitational wave in the near future is small because of the extremely small amplitudes to be expected, one is nevertheless interested in the rigorous predictions of the theory for the sake of fundamental insight.

In this chapter we will consider classical radiation fields; (no quantum aspects will be discussed). As a (local) definition of "pure" radiation we take the existence, in an empty space time V_4, of a congruence of geodetic null lines with vanishing distortion.[1] Such a congruence, the so-called "ray congruence" existes in every electromagnetic radiation field, ($\vec{E}\perp\vec{H}$, $|\vec{E}| = |\vec{H}|$). Its existence is equivalent, as has been finally shown by GOLDBERG and SACHS[2], to WEYL'S conform tensor being of special PETROV type; a property on which PIRANI[3] tried to base a local definition of pure radiation in 1957, and which

1) This definition has been succesfully introduced by J. Robinson.- A much more general notion of radiation is discussed in R. Arnowitt, S. Deser and C.W. Misner: Energy and the criteria for radiation in general relativity, preprint 1959.

2) Private communication. In the case of twistfree ray congruences the proof will be found below.

3) F.A.E. Pirani, Phys. Rev. 105, 1089 (1957)

stands in close analogy to corresponding properties of MAXWELL'S field tensor in electromagnetic radiation fields.[4)]

To be more specific, this chapter will essentially be concerned with the fields containing a non-expanding ray congruence; that is with fields which are to be expected at large distances from their sources where their rays can be regarded parallel. As for the counterpart of expanding radiation fields (with normals rays) we refer the reader to [5)]. Most of our results have already been published in [6)], but mainly without proofs. In this chapter we endeavour to close this gap by giving a detailed survey of the non-expanding radiation fields and their various subclasses.

We will start in § 2 with a short treatment of geodetic null congruences[7)], which are the important tool in pure radiation theory. In § 3 we give an approach to the study of fields containing a twistfree null congruence. § 4 will apply these tools to the class of expansion-free radiation fields in order to obtain explicit formulae for their metric forms. Some important intrinsic properties of these fields are then proven in § 5. Finally, in §§ 6 and 7, we can show a remarkable one-to-one correspondence between plane -fronted MAXWELL waves and a corresponding class of pure gravitational fields which can be

---

4) Compare our report P. Jordan, J. Ehlers and R. Sachs: "Progress in the field of general relativity", Hamburg 1960.

5) J. Robinson and A. Trautman: "Some spherical gravitational waves in general relativity", preprint Syracuse 1961.

6) W. Kundt: "The plane-fronted Gravitational Waves", Z.Phys.163 (1961).

7) For greater detail see: P. Jordan, J.Ehlers and R. Sachs: Akad. Wiss. Mainz, Abh.Math.-nat. Kl., Nr.1 (1961)

defined in the same way, and which is the simplest subclass of the expansion-free radiation fields.

§ 2. Geodetic null congruences.

A congruence of geodetic null lines or "rays" is a set of zero-length geodesics with the property that through each point of the region of space-time considered there passes one and only one element of the set. Each ray is the possible path of a test photon. A geodetic null congruence therefore corresponds to the history of a continuous cloud of photons whose velocities are regular in the sense that no collisions occur.

Formally a congruence can be described by $x^a = x^a(y^\alpha, v)$, $1 \leq a \leq 4$, $1 \leq \alpha \leq 3$, with the rays being given by $y^\alpha =$ const. For a null geodesic v can and shall be chosen such that the tangent vector $l^a \equiv \partial x^a/\partial v$ satisfies $\dot{l}_a = 0$; (here $\dot{T}_\Sigma \equiv T_{\Sigma;b} l^b$ for any tensor T_Σ). Such a parameter v is called an "affine parameter"; for each ray it is unique up to a linear transformation. The properties of a null geodetic congruence are suitably discussed by introducing a (complex) "null tetrad" $\{t_a , \bar{t}_a , l_a , m_a\}$ adapted to l_a :

$$(2,1) \qquad g_{ab} = 2\{t_{(a}\bar{t}_{b)} + l_{(a}m_{b)}\} ,$$

l_a , m_a real , $t^a\bar{t}_a = 1 = l^a m_a$, all other products zero. With respect to such a tetrad the gradient of l_a allows for the following decomposition:

$$(2.2) \qquad l_{a;b} = 2 \operatorname{Re} (z\, t_a\bar{t}_b + \sigma\, t_a t_b + \Omega\, t_a l_b + \zeta\, l_a t_b) + \beta\, l_a l_b$$

with $z = l_{a;b}\bar{t}^a t^b$, $\sigma = l_{a;b}\bar{t}^a\bar{t}^b$, $\Omega = l_{a;b} t^a m^b$

because of : $l^a l_{a;b} = 0 = l_{a;b} l^b$. In this expansion the

coefficients z, σ, ... are called "optical scalars". Putting $z = \theta + i\omega$, we call θ the "expansion", ω the "twist", $|\sigma|$ the "distortion" of the congruence, and $|\Omega|$ the "rotation" with respect to observers in the (l_a, m_a)- plane. These notations can be justified by the following considerations:

Let L be one of the rays, and L', L'' two neighbouring ones such that L', L'' meet a spacelike 2-element δF which is orthogonal to L in one of its points x, say. If δx^a is a vector connecting L and L', $\delta x^a = \frac{\partial x^a}{\partial y^\alpha} \delta c^\alpha$, it follows from the commutability of ordinary derivatives that δx^a is LIE -transferred along L:

(2.3) $$(\delta x^a)^{\cdot} = l^a{}_{;b}\, \delta x^b \quad .$$

Consequently, if $l_a \cdot \delta x^a = 0$ holds in x it holds along the whole of L, so that L', L'' meet every 2-element δF which is orthogonal on L in any one of its points. Physically this means that all photons which hit the 2-dimensional "screen" δF in the rest space of an observer in x simultaneously will also hit any other screen simultaneously which is placed orthogonal to the ray direction (in the rest space of any other observer at any later time).

We ask for the shadow cast by a small blind at $x(v)$ on a neighbouring screen $\delta F(v + dv)$ at $x(v + dv)$. Let u^a describe an observer at x, and take v such that $l^a u_a = -1$. Then

(2.4) $$h^a_b = \delta^a_b - l^a l_b + l^a u_b + u^a l_b$$

is the projection operator on the 2-element δF which is orthogonal on both l^a and u^a, for :

$$(2.5)\qquad h^a_b\, h^b_c = h^a_c\ ,\quad h^a_b u^b = h^a_b l^b = 0\ ,\quad h^a_a = 2.$$

If x^a is any vector connecting L and L', $h^a_b\, \delta x^b$ lies in δF, and its change along L is given by

$$(2.6)\qquad h^a_b\,(h^b_c\,\delta x^c)^{\cdot} = A^a_b\,\delta x^b\ ,\quad A^a_b = h^a_c\, l^c{}_{;d}\, h^d_b\,.$$

Consequently, when one advances from $x(v)$ to $x(v + dv)$, the points of intersection of neighbouring rays L' with a screen δF suffer the infinitesimal transformation:

$$(2.7)\qquad h^a_b\,\delta x^b \rightarrow (\,\delta^a_b + dv\, A^a_b)\ \ h^b_c\ \ \delta x^c$$

which is governed by the 2-tensor A^a_b. We decompose A_{ab} into a rotation ("twist"), a rotation-free similarity map ("expansion"), and a rotation-free area-conserving "distortion" according to:

$$(2.8)\qquad A_{ab} = A_{[ab]} + \tfrac{1}{2}\, A^c_c\ \ h_{ab} + (A_{(ab)} - \tfrac{1}{2}\ A^c_c\ \ h_{ab}).$$

In order to get the connection between (2.2) and (2.7) we set

$$(2.9)\qquad u^a = \tfrac{1}{2}\, l^a - m^a\ ,\ \text{so that}\ \ h_{ab} = 2\, t_{(a}\bar{t}_{b)}\ ,$$

and obtain from (2.2), (2.6) :

$$(2.10)\qquad \begin{aligned} z &= \tfrac{1}{2}\, A^c_c + i\sqrt{\tfrac{1}{2}\ A_{[ab]}A^{ab}} \\ \sigma &= \sqrt{\tfrac{1}{2}\left[\, A_{(ab)}\ A^{ab} - \tfrac{1}{2}\ (A^c_c)^2\right]}\,. \end{aligned}$$

This shows that $\omega\, dv$ describes the angle of rotation of the blind's shadow, $2\theta\, dv$ describes the relative area expansion, and $1 \pm |\sigma|\, dv$ are the eigenvalues of the (infinitesimal) distortion.

The meaning of Ω is obtained when one considers the change of the ray direction in the rest space of a geodetic observer u^a travelling through the light bundle. Again take m^a, t^a as in (2.9), (with $u^a l_a = -1$), then

$$(2.11) \qquad h^a_b \overset{\wedge}{\nabla} l^b = -2\, \mathrm{Re}(\Omega\, t^a)\ ,$$

where $\overset{\wedge}{\nabla} = u^a \nabla_a$ is the operater of covariant differentiation with respect to the observer's eigentime. This shows that $2|\Omega|$ is the magnitude of the angular velocity with which the ray direction "rotates" in the observer's comoving (parallelly propagated) rest frame.

So far our null tetrads (observers, screens) introduced in (2.1) are not unique; they can be changed by one of the following LORENTZ transformations:

1) spacelike rotation in the t^a-plane:

$$(2.12) \qquad t'^a = C\, t^a\ , \qquad C\bar{C} = 1\ ,$$

2) timelike rotation in the (l^a, m^a)-plane:

$$(2.13) \qquad l'^a = A l^a\ ,\ m'^a = A^{-1} m^a,\ \text{with } \dot{A} = 0\ ,$$

3) lightlike rotation:

(2.14) $$t'^a = t^a - \bar{B}\, l^a\ ,\quad l'^a = l^a\ ,\quad m'^a = \bar{B}\, \bar{t}^a + B\, t^a - B\bar{B}\, l^a + m^a$$

All these transformations leave the differentials $z dv$, $|\sigma| dv$ invariant. But σ can be made real along L by means of (2.12), (transformation to "principal axes of distortion"), Ω can be transformed away for $|z| \neq |\sigma|$ by means of (2.14), and ζ and β can be pointwise cancelled by means of (2.13) (with nonconstant A). It has to be kept in mind, therefore, that the optical scalars are in general no invariants of the congruence (though "almost" invariants). However, $|\Omega|$ is an invariant for $z = 0 = \sigma$.

A null congruence is "normal"(l_a is hypersurface-orthogonal) if and only if $\omega = 0$. In that case l_a may and shall be chosen to be a gradient, $l_a = u_{,a}$ (which implies $\dot{l}_a = 0$). We call u a "phase"or "retarded time" of the(normal) congruence.

A non-expanding geodetic null congruence with $R_{ab}\, l^a l^b = 0$ is twistfree (normal) if and only if it is free of distortion. In order to prove this we differentiate the defining equation for z in ray direction, where we can and shall assume the null tetrad to be parallelly propagated along L. With the help of RICCI'S identity for l_a: $l_{a;[bc]} = -\frac{1}{2} R_{dabc}\, l^d$ we then obtain:

$$\begin{aligned} \dot{z} &= l_{a;bc}\, \bar{t}^a t^b l^c \\ &= (l_{a;cb} - R_{dabc}\, l^d)\, \bar{t}^a t^b l^c \\ &= - l_{a;c}\, l^c{}_{;b}\, \bar{t}^a t^b + \tfrac{1}{2} R_{cd}\, l^c l^d \end{aligned}$$

(2.15) $$= -(z^2 + \sigma\bar{\sigma}) + \tfrac{1}{2} R_{cd}\, l^c l^d \quad .$$

Under the above assumptions the real part of this equation reduces to $\omega^2 = |\sigma|^2$, and our statement is evident.

A {spacelike} 2-surface orthogonal on l_a is called "Wave surface".. We remark that in this notation the spheres (t,r) = const of SCHWARZSCHILD's static vacuum fields are (frozen) wave surfaces. We prove:

Lemma 2.1 Necessary and sufficient for the existence of (finite) wave surfaces in a geodetic null congruence is the vanishing of its twist.

Proof: Suppose there exist wave surfaces spanned by t^a (that is: spanned by $Re(t^a)$, $Im(t^a)$). The surface-forming character of t^a is described by

$$\overset{t}{\nabla}\bar{t}^a - \overset{\bar{t}}{\nabla} t^a = Im(\alpha t^a), \quad \overset{t}{\nabla} \equiv t^a \nabla_a, \quad \text{so that:}$$

$$0 = l_a(\overset{t}{\nabla}\bar{t}^a - \overset{\bar{t}}{\nabla} t^a) = 2 l_{a;b}\, t^{[a}\bar{t}^{b]} = -2i\omega .$$

Conversely, the explicit shape of the metrics (3.2) below will tell that the surfaces (S, u) = const are wave surfaces; we will come back to this point.

§ 3 Radiation metrics

An important question in general relativity theory is the following: how does the metric field$ behave at large distances

from fast moving sources? From characteristic theory one knows that disturbances of a static region (signals) spread with the velocity of light, (perhaps leaving a wake behind them). In the study of wave fields it is therefore reasonable to adapt the coordinate system to the wave fronts, that is to the null hypersurfaces of equal phase (if existing). In other words: it is reasonable to study metrics possessing a normal null congruence (which is automatically geodetic, see above). Now it is well-known [8] that locally a normal-hyperbolic V_4 always contains such a congruence. That is there always exist local solutions of the second ~~order~~ degree differential equation $g^{ab} u_{,a} u_{,b} = 0$; whereby "local" means that there exist coordinate balls around each point of an initial hypersurface into which a solution can be continued. But it is an explicit (though weak) assumption when we consider coordinate cubes with this property.

Let $l_a = u_{,a}$ be the tangent vector of a twistfree null congruence in some region of space-time (V_4 of signature + 2) $u^{,a} u_{,a} = 0 = \dot{l}_a$. We use two (necessarily spacelike) coordinates x^A , (A,B,... = 1,2), whose gradients are linearly independent

---

8) See e.g. V.Fock: Theorie von Raum,Zeit und Gravitation, Akademie Verlag Berlin 1960, Anhang D.

of l_a but orthogonal on l_a : $x^A{}_{,b} l^b = 0$; they exist locally according to a well known existence theorem on linear (homogeneous) differential equations. If s is a fourth coordinate with $s_{,b} l^b \neq 0$, we have an independent system $(x^A, s, u) \doteq (x^a)$, $1 \leq a \leq 4$, with:

$$(3.1) \qquad g^{44} = 0 = g^{A4} ,$$

whence $g_{3a} = 0$ for $a \neq 4$ because of $g^{4a} g_{ab} = \delta^4_b$.
That is our metric form G can be written in the form:

$$(3.2) \qquad G = g_{AB}(dx^A + B^A du)(dx^B + B^B du) + 2e^C ds du + H du^2$$

with arbitrary (real) functions g_{AB}, B^A, C, H of all coordinates satisfying $\det (g_{AB}) > 0$ (in order to guarantee signature $+2$).

At this stage we recognize that the 2-surfaces $(s,u) = \text{const}$ are spacelike and orthogonal to l_a , hence "wave surfaces".

We intend to study our metric (3.2) in relation with the properties of its preferred null congruence. From (2.10) we find by direct calculation:

$$(3.3) \qquad \theta = \frac{1}{4} e^{-C} \partial_3 \ln \det (g_{AB}) = \frac{\dot{r}}{r} , \quad (\dot{} = \dot{\nabla}),$$

(3.4) with $\gamma \equiv \sqrt[4]{\det \; (g_{AB})}$.

<u>1)</u> Suppose $\theta \neq 0$. In this case γ is called a <u>ratio-of-area distance</u> according to the proportionality:

$dr/r = \theta\, dv$; (screen areas are proportional to r^2). We write $g_{AB} = r^2 h_{AB}$ and call h_{AB} the <u>"ray metric"</u> ; the determinant of h_{AB} is independent of s . It is now useful to introduce r instead of s as the third coordinate ; (because of $r,_a l^a = r\theta \neq 0$ this is admissible). From (3.3) we then get $\theta\, e^C{}_r = 1$, so that (3.2) becomes:

$$(3.5) \qquad G = r^2 h_{AB}(dx^A + B^A du)(dx^B + B^B du) + \frac{2drdu}{\theta r} + H du^2 ,$$

$$\text{with } \partial_r \det (h_{AB}) = 0.$$

Next we calculate σ , using the definition (2.2):

$$(3.6) \qquad \sigma = \frac{\theta r}{2} g_{ab,3} \bar{t}^a \bar{t}^b$$
$$= \frac{\theta r^3}{2} \bar{t}^A \bar{t}^b \partial_r h_{AB} .$$

This shows that σ vanishes if and only if h_{AB} is independent of r , (compare (3.5)).

The case of vanishing σ, that is of a <u>distortion-free</u> geodetic null congruence, is of outstanding importance: it has been mentioned in the introduction that such a "ray congruence" exists in every

electromagnetic radiation field, and will be taken as the defining property for a gravitational radiation field. For $\sigma = 0$ (3.6) implies that we can introduce isometric coordinates in the (2-dimensional!) wave surfaces $(r,u) = \text{const}$ by means of a transformation: $x'^A = f^A(x^B, u), (x'^a = x^a (\text{otherwise}))$, so that we have:

$$(3.7) \qquad h_{AB} = p^2 \delta_{AB}, \quad p,_r = 0 .$$

If furthermore $R_{cd} l^c l^d = 0$ is assumed, the real part of (2.15) gives $\dot{\theta} = -\theta^2$ so that

$$(3.8) \qquad \theta r = 1$$

can be achieved by suitable choice of the origin of r.
The equations $R_{ab} t^a t^b = 0$ next imply that by another "gauge transformation" we can arrive at $B^A = 0$, so that the metric form of an expanding <u>pure radiation field</u> with $R_{ab} l^a l^b = 0 = R_{ab} t^a t^b$ can be brought to the shape

$$(3.9) \qquad G = (r p)^2 (dx^2 + dy^2) + 2\, dr du + H\, du^2 .$$

We stop here, because a thorough treatment of these fields can be found in [5]; and turn to the case of non-expanding radiation fields.

2) Suppose $\theta = 0$. Here we get from (3.3) :

$$\partial_r \det (g_{AB}) = 0 ,$$

and $\int e^{C} ds = v$ can be introduced as coordinate instead of s ; v is an affine parameter, for we have $v_{,a} l^a = 1$. From (2.15) we learn that the additional assumption $R_{ab} l^a l^b = 0$ is equivalent with $\sigma = 0$. This time we find:

$$\sigma = \frac{1}{2} g_{ab,3} \bar{t}^a \bar{t}^b \quad , \tag{3.10}$$

so that again $\sigma = 0$ is the necessary and sufficient condition on the ray metric g_{AB} to be transformable into the isometric form

$$g_{AB} = p^2 \delta_{AB} , \quad p_{,v} = 0. \tag{3.11}$$

That is we have for the metric of the general non-expanding pure radiation field :

$$G = p^2 \left\{ (dx + B^1 du)^2 + (dy + B^2 du)^2 \right\} + 2dvdu + Hdu^2 \tag{3,12}$$
with $p_{,v} = 0$.

The rest of this chapter is devoted to an analysis of this class of fields.

§ 4. The expansion-free radiation fields.

Up to now we have defined "pure radiation" fields by the existence of a ray congruence without being specific about field equations. From now on let us restrict considerations to combined EINSTEIN-MAXWELL -fields in empty space, whereby the electromagnetic field is assumed to be null ($\vec{E} \perp \vec{H}$, $|\vec{E}| = |\vec{H}|$). That is we deal with the equations:

$$(4.1)\quad \left\{\begin{array}{l} -R_{ab} = T_{ab} = \frac{1}{2}\Phi_{ac}\bar{\Phi}_b{}^c \ , \\ \Phi^{ab}{}_{;b} = 0 = \Phi_{ab}\Phi^{ab} \ , \end{array}\right\}$$

where $R_{ab} \equiv R^c{}_{acb}$ is the RICCI-tensor, T_{ab} the energy momentum tensor, $\Phi_{ab} \equiv F_{ab} + i\,\overset{*}{F}_{ab}$ the (complex) MAXWELL-bi tor, and the star stands for the (real) dual operation: $\overset{*}{F}_{ab} \equiv \frac{1}{2}$; F^{cd}, $\overset{**}{F}_{ab} = -F_{ab}$. The (last) equation stating the nullity of F_{ab} is equivalent with:

$$(4.2)\qquad \Phi_{ab} = 2\,k_{[a}t_{b]} \ , \quad k^a k_a = 0 = k^a t_b ,$$

t_a complex as always. Proof: (4.2) implies $\Phi_{ab}\Phi^{ab} = 0$. Conversely we have $\frac{1}{2}\Phi_{ab}\Phi^{ab} = F_{ab}(F^{ab} + i\,\overset{*}{F}{}^{ab})$; the vanishi of $F_{ab}\overset{*}{F}{}^{ab}$ is equivalent with F_{ab} being the skew product of two vectors[9]; $F_{ab}F^{ab} = 0$ then implies that these two vectors span a lightlike 2-surface (for $\Phi_{ab} \neq 0$, otherwise the statement is trivial).

Inserting (4.2) into (4.1) our field equations reduce to:

$$(4.3)\qquad \left\{\begin{array}{l} 2R_{ab} = -k_a k_b \\ \Phi^{ab}{}_{;b} = 0 \ , \ \Phi_{ab} \text{ as in (4.2).} \end{array}\right\}$$

---

9) See e.g.: P. Jordan, J. Ehlers, W. Kundt, Akad. Wiss. Mainz, Abh. Math.-Nat. Kl. 1960, Nr. 2; Abschnitt 1.2.

We do not exclude the case of pure gravitational fields for which ϕ_{ab} vanishes, and (4.3) simplify to: $R_{ab} = 0$. If ϕ_{ab} does not vanish, however, MAXWELL's equations $\phi^{ab}{}_{;b} = 0$ imply that the k_a - congruence is a ray congruence (= distortion-free geodetic null congruence). Namely as follows:

$$\phi^{ab}{}_{;b} = k^a t^b{}_{;b} + \overset{t}{\nabla} k^a - t^a k^b{}_{;b} - \overset{k}{\nabla} t^a , \tag{4.4}$$

so that $k_a \phi^{ab}{}_{;b} = -k_a \overset{k}{\nabla} t^a = t_a \overset{k}{\nabla} k^a$. From (4.3) and $0 = k_a \overset{k}{\nabla} k^a$ ~~a k~~ we therefore get: $\overset{k}{\nabla} k^a \sim k^a$, which means that k^a is geodetic. We write:

$$k_a = \varkappa \, l_a \quad \text{with} \quad \dot{l}_a = 0 . \tag{4.5}$$

for suitable scalar $\varkappa$. Next we get from (4.3) and $t_a \cdot (4.4)$:

$$0 = t_a \overset{t}{\nabla} k^a = \varkappa \, l_{a;b} t^a t^b = \varkappa \bar{\sigma} ,$$

proving $\sigma = 0$.

On the other hand it can be shown that for every ray congruence k_a there exists a MAXWELL bivector ϕ_{ab} satisfying (4.3) with the possible exception of a positive (non-constant) scalar factor λ in the EINSTEIN equation: $2\lambda R_{ab} = - k_a k_b$. A first proof of this theorem was given by I. ROBINSON [10], a more systematic treatment can be found in [11]. The idea of proof is as follows:

10) I. Robinson: Journ. of Math. Phys. 2, No 3, p.290 (1961)

11) P. Jordan, J. Ehlers, R. Sachs, Akad. Wiss. Mainz, Abh. math.-nat. Kl. Nr. 1, 1961, Abschnitt 3.3.

Φ_{ab} must necessarily be of the form

(4.6) $$\Phi_{ab} = 2e^{i\varphi} k_{[a} t_{b]} , \ \varphi \text{ real} , \ t^{a} \bar{t}_{a} = 1 ;$$

it can then be shown that for a geodetic and distortion-free k_a there exists a scalar φ such that (4.6) satisfies $\Phi^{ab}{}_{;b} = 0$. φ is for $\omega \neq 0$ unique up to an additive constant.

The equation $\lambda = 1$, equivalent with

(4.7) $$(R_{ab} + T_{ab}) m^a m^b = 0,$$

is nasty. It is subject of a recent paper by W. KUNDT [12] where it is reformulated as a condition on the metric and its derivatives alone. In what follows we will be lax about this one equation; that is we will ignore it except in simple cases, as it does not provide one with deeper insight into intrinsic properties.

We are now ready to present a reasonable definition of expansion-free radiation fields:

Definition 4.1: A space-time (V_4 of signature $+2$) is called an expansion-free radiation field if it contains a non-expanding ray congruence l_a, and if it satisfies the field equations

---

12) P. Jordan, W. Kundt: Akad. Wiss. Mainz, Abh. math.nat. Kl. Nr. 3, 1961

$$(4.8) \qquad 2\, R_{ab} = -\mu l_a l_b\,, \quad \mu \geq 0,$$

together with (4.7).

The rest of this paragraph is devoted to a derivation of simple <u>metric forms</u> for this class. In (3.12) we already arrived at

$$G = p^2 \left\{ (dx + B^1 du)^2 + (dy + B^2 du)^2 \right\} + 2dvdu + Hdu^2$$
$$\text{with} \quad p_{,v} = 0\,,$$

where use was made of $R_{ab} l^a l^b = 0$, and where $l_a = u_{,a}$, (a non-expanding ray congruence with $R_{ab} l^a l^b = 0$ is normal!). We are now left to satisfy the rest of (4.8), and possibly (4.7).

Before so doing it is convenient to introduce further complex notation by: $x + iy \doteq z$, $B^1 + i B^2 \doteq B$, so that G becomes:

$$(4.9) \qquad G = p^2 \left| dz + Bdu \right|^2 + 2\, dvdu + Hdu^2$$
$$\text{with} \quad p_{,v} = 0.$$

In (4.9) the coordinates (z,v,u) are unique up to the following <u>gauge transformations:</u>

$$(4.10) \left\{ \begin{array}{ll} \text{I} & \bar{u} = f(u),\ \bar{v} = f'^{-1}(u)\, v,\ \bar{z} = z \text{ (change of phase)} \\ \text{II}. & \bar{v} = v + f(x,y,u),\ \bar{z} = z,\ \bar{u} = u. \text{ (change of affine origin)} \\ \text{III}. & \bar{z} = F(z;u),\ F \text{ analytic in } z, \bar{v} = v, \bar{u} = u \text{ (conformal} \\ & \qquad\qquad \text{change in wave surfaces).} \end{array} \right\}$$

These gauges will be helpful in further simplifying the fundamental form.

Now the equations $0 = R_{ab} l^a t^b$ amount to $B_{,33} = 0$.

Two cases have to be distinguished:

1) $B_{,3} = 0$, which is equivalent with $\Omega = 0$ as will be shown in the next paragraph. In this case the equations $0 = R_{ab} t^a t^b$ are empty, and from $0 = R_{ab} t^a \bar{t}^b$ we get $\Delta \ln p = 0$. This last equation is equivalent to p being the magnitude of an analytic function of z (still depending on u) so that a gauge III can be applied (with $|F'| = p^{-2}$) to arrive at $\bar{p} = 1$. [13] Next we can make B real using II. The rest of (4.8) then leads to

$$G = |dz + b\,du|^2 + 2\,dv\,du + H\,du^2 \quad \text{with}$$

$$(4.11)\quad \left\{ \begin{array}{l} b \text{ real}, \; b_v = 0 = \Delta b, \; (b_a \equiv \partial_a b, \; \Delta \equiv \partial_x^2 + \partial_y^2) \\ \quad H = -v\,b_x + A, \quad A_v = 0, \\ \Delta A + (2b\partial_x + 3b_x - 2\partial_u)\,b_x + b_y^2 = \Lambda \end{array} \right\}$$

It is of interest to know if further simplifications of (4.11) can be achieved by means of a gauge (4.10). The answer is: If and only if b is linear in x and y:

$$(4.12)\qquad b_{,AB} = 0 ,$$

---

13) On the other hand we could have used $0 = \Delta \ln p$ to make $B = 0$, at the expense of having $p = 1$.

b can be transformed away by means of II and III (with $F = e^{i\phi(u)} + g(u)$, $f = a(u)x^2 + b(u)xy$, $g(u)$ complex, all other functions real). These latter fields $(b = 0)$ are the so-called "plane-fronted waves with parallel rays" (= pp waves), compare [9),14)]. They have been independently discovered by [15)] and by J. ROBINSON (1956).

2) $B_{,3} \neq 0$, $(\Leftrightarrow \Omega \neq 0)$. Because of $B_{,33} = 0$ we have $B \equiv vD + C$ with $D_{,v} = 0 = C_{,v}$, (D,C complex). Now the equations $0 = R_{ab}\, t^a t^b$ read:

(4.12) $$D_{,1} + i\,D_{,2} = -\frac{1}{2}\,p^2 D^2$$

with $D \neq 0$ by assumption. Putting $\phi D^{-1} \equiv b^1 - ib^2$ we get $$b^1{}_{,2} - b^2{}_{,1} = 0, \quad b^1{}_{,1} + b^2{}_{,2} = -\frac{1}{2}p^2,$$
whence $$b^A = b_{,A}, \quad \Delta b = -\frac{1}{2}\,p^2 .$$

---

14) Ehlers, J., and W. Kundt: Chap. 3 in: The theory of gravitation, edit. by L. Witten. To appear at Wiley, New York.

15) Brinkmann, H.W.: Proc. Nat. Acad. Sci. Wash. 9, 1 (1923). - Hely, J.: C.R. Acad. Sci. Paris 249, 1867 (1959). - Peres, A.: Phys. Rev. Letters 3, 571 (1959).

The equation $R_{ab}\, t^a \bar{t}^b = 0$ gives a second relation between b and p ; for later use we want to present it in two different notations, each being suited in its own way. We rewrite the general radiation metric (3.2) :

$$(4.13) \qquad G = g_{AB} dx^A dx^B + 2\, m_a dx^a du ,$$

$$\text{where} \quad m_a \equiv \left(g_{AB} B^B ,\; e^C ,\; \tfrac{1}{2}\left[H + B_A B^A\right]\right),\; \det(g_{AB}) > 0.$$

(Note: m_a is the same as in (2.1)!). In our present case (3.12) we have $g_{AB} = p^2 \delta_{AB}$, $e^C = 1$, so that:

$$(4.14) \qquad m_a = \left(p^2 B^1 ,\; p^2 B^2 ,\; 1,\; \tfrac{1}{2}\left[H + p^2 |B|^2\right]\right).$$

Now $R_{ab} t^a \bar{t}^b = 0$ takes the form:

$$(4.15) \qquad 2\Delta \ln p = m_{1,31} + m_{2,32} - \tfrac{1}{2}\left({m_{1,3}}^2 + {m_{2,3}}^2\right), \text{ with}$$

$$(4.15') \qquad m_A = p^2 \left({b_{,1}}^2 + {b_{,2}}^2\right)^{-1} b_{,A} \;.$$

If (4.15') is inserted in (4.15) , and the ray metric $p^2 \delta_{AB}$ is used to raise capital latin indices (which in our applications will be the same as using the full matric g_{ab}) we can collect our results into:

$$G = p^2 \left|dz + B du\right|^2 + 2 dv du + H du^2, \text{ with}$$

$$(4.16) \quad \left\{ \begin{array}{l} B = v\left(p^2 b_{,A} b^{,A}\right)^{-1}\left(b_{,1} + i\, b_{,2}\right) + C \;, \quad p, b, H \text{ real} , \\ p_{,v} = 0 = b_{,v} = C_{,v} \;, \quad \partial^A_A b = -\tfrac{1}{2} ,\; \partial^A_A \equiv g^{AB} \partial_B \partial_A ,\; C \text{ compl} \\ b^{,A} b^{,B} b_{,AB} = b_{,A} b^{,A} \left[b^{,B} \partial_B \ln p - \tfrac{1}{2}\right] - \left(b_{,A} b^{,A}\right)^2 \partial^B_B \ln p \end{array} \right\}$$

for the general expansion-free radiation field with non-vanishing rotation $|\Omega|$. The 3 equations $0 = R_{ab}\, t^a m^b = R_{ab}\, l^a m^b$ then determine $H_{,v}$ up to an additive function of u which can be cancelled with (4.10) I. H becomes a polynomial of second degree in v.

Equations (4.16) look rather complicated, hence it is of interest to find special solutions. If we assume $b_{,2} = 0$, (which means, at least for $H_{,2} = 0$, that there exists a group G_1 of isometries whose orbits lie in wave surfaces), equations (4.16) simplify to $(b' \equiv b_{,1})$:

$$(4.17) \quad \left\{ \begin{array}{ll} a) & b'' = -\frac{1}{2} p^2 \\ b) & p^2 (\ln p)' = b' (\ln p)'' \,. \end{array} \right\}$$

For $p' = 0$ the general solution of (4.17) can be reduced to

$$(4.18) \qquad p = 1, \quad b = -\frac{x^2}{4}$$

by means of (4.10) III, (u-dependent translation of x), and because only $b_{,A}$ enters into the metric. This is a class of type III fields and will be considered below. For $p' \neq 0$ we get from (4.17):

$$b'^{-2} = 4c_1 (\ln p)'$$

where from now on $c_i = c_i(u)$ are "constants" of integration.

Multiplying by (4.17) a), we find, after a second integration:

$$b'^{-1} = c_1 p^2 + c_2 .$$

Again using (4.17) a) and integrating we get

$$c_1 (b')^2 - c_2 b = x + c_3 , \tag{4.19}$$

the general solution of which can only be given in implicit form. An elementary solution is obtained for $c_2 = 0$, it can be reduced to:

$$p = x^{-1/4} , \quad b = -\frac{2}{3} x^{3/2} . \tag{4.20}$$

In (4.18) we found an integral of (4.16) with <u>$\Delta \ln p = 0$.</u> This class of metrics will turn out below to be of special interest. Here we want to show that (4.16) is already the <u>general solution</u> with this property (after suitable coordinate gauges): As mentioned in the derivation of (4.11), $\Delta \ln p = 0$ is the necessary and sufficient condition for "$p = 1$ after suitable gauge". We therefore assume $p = 1$. Now the equations (4.12), (4.15) can be written as follows:

$$\left\{ \begin{array}{l} 2D^1{}_{,1} = (D^1)^2 , \quad 2D^2{}_{,2} = (D^2)^2 \\ D^1{}_{,2} + D^2{}_{,1} = D^1 D^2 \end{array} \right\} \tag{4.21}$$

where $\qquad D^1 + iD^2 \equiv D = m_{1,3} + i\, m_{2,3} .$

As we are still dealing with the case $B_{,3} \neq 0$, (the case $B_{,3} = 0$ was fully overcome in (4.11)), we have $D \neq 0$. If $D^2 = 0$ we are back at our earlier assumption $b_{,2} = 0$ which was mastered in (4.18). We therefore assume $D^1 D^2 \neq 0$, and intend to show that it can be reduced to (4.18): From the first line of (4.21) we get, by direct integration:

$$(4.22) \qquad D^1 = -2(x + f(y;u))^{-1} , \quad D^2 = -2(y + g(x;u))^{-1} .$$

Now the second line reads

$$(4.23) \qquad f_y (y + g)^2 + g_x (x + f)^2 = 2(x + f)(y + g),$$

with $f_y \equiv \partial_y f$, ... Partial differentiation with respect to x and y and linear combination yields:

$$(4.24) \qquad \left\{ \begin{array}{l} g_{xx} (x + f)^3 = f_{yy} (y + g)^3 \\ f_{yy} (y + g)^2 = 2(x + f)(1 - f_y g_x) \end{array} \right\} .$$

Assume $f_{yy} \neq 0$. In this case the first line can be rewritten as

$$(f_{yy})^{-1/3} (x + f) = (g_{xx})^{-1/3} (y + g) ,$$

and differentiation with respect to x, loving consideration, and integration leads to

$$f = c\,y^{-1} , \quad g = c\,x^{-1} , \quad c = c(u) ,$$

a solution which violates the second line of (4.24). We therefore have: $f_{yy} = 0 = g_{xx}$, and: $f_y g_x = 1$; or:

$f = c_1 y + c_2 \quad , \; g = c_1^{-1} x + c_3 \, , \quad c_i = c_i \,(u).$

Insertion into (4.23) gives $c_1 c_3 = c_2$, so that finally (see (4.22)):

$$D = -\,2\, e^{i\varphi} \,(x \cos\varphi + y \sin\varphi + c)^{-1} \,, \tag{4.25}$$

~~and~~ where $\cos\varphi = (1 + c_1^{\,2})^{-1/2}$. We now apply a gauge (4.10) III with $F = z\, e^{-i\varphi} + c$, and arrive at

$$D = -\frac{2}{x} \quad , \tag{4.26}$$

which agrees with (4.18) , and finishes our proof.

As a final effort of this paragraph we want to evaluate the still outstanding field equations $\underline{2 R_{ab}\, m^a = -\mu l_b}$ for the class of metrics $\underline{(4.16),\ (4.18)\ :}$ In the notation of (4.14), $0 = R_{ab}\, m^a l^b$ reads (in general):

$$2\, m_{4,33} = p^{-2} \left\{ m_{1,3}^{\;2} + m_{2,3}^{\;2} - m_{1,31} - m_{2,32} \right\} , \tag{4.27}$$

so that for $p = 1$ we get from (4.15) :

$$m_{4,33} = \frac{1}{4} \left| B_{,3} \right|^2 \qquad ,$$

or:

$$m_{4,3} = \frac{v}{4} \left| D \right|^2 + F\,(x^A, u) \ . \tag{4.28}$$

Next we specialize D to (4.26), and use a gauge II (with $\partial_y f = \mathrm{Im}(C)$) to make C in (4.16) real. The equations $0 = R_{ab}\, m^a t^b$ are linear in v; because of (4.21), however, only the v-independent terms survive; they shrink, after these

simplifications, to

$$(4.29)\quad \left\{\begin{array}{ll} F_{,x} = (\tfrac{1}{2}\partial_y^{\,2} + x^{-2})\, C & ,\ \Delta = \partial_x^{\,2} + \partial_y^{\,2}, \\ F_{,y} = -(\tfrac{1}{2}\partial_x + x^{-1})\, C_{,y} & \quad C \text{ real}\,, \end{array}\right\}$$

with the integrability condition: $0 = \partial_y \Delta x C$. This latter condition guarantees that by means of another gauge II (with $\partial_y f = 0$) we can effect

$$(4.30)\qquad \Delta x C = 0\ ,$$

and (4.29) can be integrated in closed form:

$$(4.31)\qquad F = -(x^{-1} + \tfrac{1}{2}\partial_x)\, C\ .$$

Here we have ignored an additive function of u as it can be gauged away by means of (4.10) I.

The last equation $2 R_{ab} m^a m^b = -\mu$ is of second degree in v, but for the class of metrics under consideration only the *zeroth* degree part survives. We present it in the following list of results: the general <u>non-recurrent type III</u> ($\Leftrightarrow B_{,3} \neq 0,\ p = 1$) <u>expansion-free</u> radiation metric can be brought to the shape:

$$(4.32)\quad \begin{array}{l} G = |dz - (2 v x^{-1} + c)\, du|^2 + 2\, dv du + H du^2 \\ \left\{\begin{array}{l} c,\ H \text{ real}\ ,\quad c_v = 0 = \Delta x c\ , \\ H = -3 v^2 x^{-2} + v\,(\partial_x - 2_x^{-1})\, c + A,\ A_v = 0\ , \\ x \Delta x^{-1} A + 2c\,(\partial_x - x^{-1})\, c_x + 3\, c_x^{\,2} + c_y^{\,2} + 2 c_{xu} = \mu \end{array}\right\}, \end{array}$$

where we have put $c = -C$.

How can the remaining gauges (4.10) be used? A gauge II leads to $\bar{c} = 0$ if

$$(2 - x\partial_x)f = xc$$

whereby the conservation of the form of (4.32) can be achieved by III with $F = z + i\,g(u)$ if and only if $\partial_A f_y = 0$. The latter condition implies

$$\partial_A \partial_y xc = 0 \,, \tag{4.33}$$

so that xc must depend linearly on x^A; (on account of (4.32), (4.33) can be replaced by $\partial_B \partial_A xc = 0$). But then

$$f = \frac{1}{2}(xc + x\partial_x xc)$$

really effects $\bar{c} = 0$, and (4.33) is the necessary and sufficient condition on c to be superfluous.

§ 5. Properties of the expansion-free radiation fields.

In order not to crush the reader to death with non-ending calculations when dealing with the properties of the expansion-free radiation fields, we have tried to anticipate most of the formal treatment in the preceding paragraph. We are now going to fill the formulae obtained above by intrinsic meaning. Let us start with

Theorem 5.1 : 1) WEYL's conform tensor of an expansion-free radiation field is of special PETROV-type. 2) The ray congruence of an expansion-free radiation field is unique except possibly

in the case of a D-field.

<u>Remark</u>: In this theorem the adjective "expansion-free" can be replaced by "pure", compare our remark in § 1 ; for a proof see [5].

Proof of theorem 5.1: In what follows we will make use of some <u>conform tensor</u> properties, and of the notations presented in our earlier report [4], or in [7]. Especially BEL's type criteria ([4], (II.25), page 30) by means of multiple principal null directions and the ordering of types proposed by PENROSE ([4], diagram page 26) will be taken to be known. With this we can determine the type of our radiation metrics by means of RICCI's identity

$$(5.1) \qquad -2\, l_{a;[bc]} = l^d\, C_{dabc} \; ,$$

in which we could replace the full RIEMANN tensor R_{dabc} by WEYL's (tracefree dual symmetric) conform tensor. C_{dabc} in virtue of our field equations (4.8) ; namely according to the general decomposition formula:

$$(5.2) \qquad \left\{ \begin{array}{l} R_{abcd} = C_{abcd} + \frac{R}{12}\, g_{abcd} - g_{abe[c} S^e_{d]} \; , \\ g_{abcd} \equiv 2 g_{a[c} g_{d]b} \; , \; S_{ab} \equiv R_{ab} - \frac{R}{4}\, g_{ab} \; , \end{array} \right\}$$

and because of $R = 0 = S^a_b\, l^b = S^a_{[b}\, l_{c]}$ for pure radiation metrics. C_{abcd} in (5.1) is special if and only if

$$(5.3) \qquad l^b\, l_{[d}\, l_{a];[bc]} = 0 \; ;$$

it is of type III or more special if and only if

(5.4) $$l_{[d}\, l_{a]\,;[bc]} = 0 \quad ,$$

and of type N or O if and only if

(5.5) $$l_{a;[bc]} = 0 \quad .$$

Now we get, for our general non-expanding radiation metric (4.9) respectively (4.13), (4.14):

(5.6) $$l_{a;b} = \frac{1}{2}\, g_{ab,3} = l_{(a}\tau_{b)} + \beta\, l_a l_b \quad , \text{ with}$$

(5.7) $$\left\{ \begin{array}{l} \tau_a = 4\,\mathrm{Re}(\Omega t_a) = \delta_a^B\, m_{B,3}, \ (\text{or: } \tau_1 + i\tau_2 = p^2 B_{,3}) , \\ |\Omega| = \frac{1}{p\sqrt{8}}\,(m_{1,3}^{\;2} + m_{2,3}^{\;2})^{1/2} = \frac{p}{\sqrt{8}}\,|B_{,3}| , \\ \beta = m_{4,3} = \frac{1}{2}\,\partial_3 (H + p^2 |B|^2) \ ; \ \tau_a l^a = 0 \quad ; \end{array} \right\}$$

and $B_{,33} = 0$ implies

5.8) $$l_{[a}\,\dot{\tau}_{b]} = 0.$$

From (5.6) we obtain

(5.9) $$-2\, l_{a;[bc]} = l_a (\tau_{[c,b]} + 2\, l_{[c}\, \beta_{,b]}) + \tau_{a;[b}\, l_{c]} + \frac{1}{2}\,\tau_a\, l_{[b}\tau_{c]} ,$$

proving (5.3) to be satisfied, so that 1) of the theorem is proven.

In order to admit 2) of the theorem we remember that a multiple principal null direction l^a (defined by (5.3)) is unique except for type D (of the conform tensor). In the latter case there exist exactly two congruences l^a which are both geodetic and distortion-free; it is a still open question whether "the second congruence" is always normal , or perhaps even

expansion-free. Examples will be given below where it is expansion-free, see (5.17).

Next we ask for all metrics which are of type III or more special. They can be characterized by

Theorem 5.2: The fields of type III, N, or 0 are characterized as the only expansion-free radiation fields which possess plane wave surfaces. Formally they are distinguished as the fields for which "$p = 1$" can be achieved. They were obtained in (4.11), ($\Omega = 0$), and (4.32), ($\Omega \neq 0$).

Remark: A corresponding theorem for the class of pure radiation fields with a normal [16], expanding ray congruence reads as follows: In this latter class the only fields whose preferred[17] wave surfaces (r, u) = const have constant GAUSSian curvature are the N-fields, together with the SCHWARZSCHILD-like D-fields. (By the "SCHWARZSCHILD-like" vacuum fields we mean 1) SCHWARZSCHILD's one-parametric spherically symmetric class, 2) a similar one-parametric class obtained from the first by replacing the positive curvature wave surfaces by such of constant negative curvature, 3) a limiting case of both, with plane wave surfaces. They were called the

---

16) Remember lemma 2.1: wave surfaces exist only in normal (null) congruences.

17) They are the surfaces of constant "distance" from the "source" at $r = 0$.

"A-fields" in [18] where one finds a complete list of the degenerate static vacuum fields). Only the A3-field, mentioned in parantheses, has <u>plane wave surfaces</u>. A proof of this remark can be collected from [5].

Proof of theorem 5.2: The condition (5.4) for type III (or a more special type) implies the weaker condition:

(5.10) $$0 = \tau^c \tau^a l_{a;[bc]} \hat{=} l_b J ,$$

and for $\Omega \neq 0$ we find from (5.7), (4.15'), and the last line of (4.16):

$$J = \frac{1}{4} t^4 [t^{-2} \tau^a \partial_a \ln t - \tfrac{1}{2}] \text{ with } t^2 \hat{=} \tau^a \tau_a = p^2 (b_{,1}{}^2 + b_{,2}{}^2)^{-1},$$

$$\Rightarrow J = \frac{1}{4} t^4 [b^{,A} \partial_A \ln t - \tfrac{1}{2}]$$

(5.11) $$= (4 p^2 b_{,A} b^{,A})^{-1} \Delta \ln p .$$

If $\Omega = 0$, (5.10) holds identically true; if $\Omega \neq 0$, (5.11) shows that $\Delta \ln p = 0$ is a necessary condition for type III to occur. But we have shown in § 4 that all metrics satisfying $\Delta \ln p = 0$ can be transformed into (4.32).

Now we find from (5.7) and respectively (4.11), (4.32):

(5.12) $$\left\{ \begin{array}{l} \Omega = 0, \quad \beta = -\frac{1}{2} b_{,x} \\ -2\, l_{a;[bc]} = l_a l_{[b}\, b_{,c]x} \end{array} \right\}$$

for the metrics (4.11) ; and after a lengthy calculation:

---

18) J. Ehlers and W. Kundt, Chap 2 in: The theory of gravitation, edited by L. Witten, to appear at Wiley, New York.

$$(5.13)\qquad \left\{ \begin{array}{l} |\Omega| = (\sqrt{2}\,x)^{-1} \quad , \quad \tau_a = -2\,\partial_a \ln x \ , \\ -2\, l_{a;[bc]} = l_a\, l_{[b}\chi_{c]} \ , \\ \chi_a = \frac{1}{2}\tau_a\, \partial_x{}^2 (xc) - \delta_a^2\, x^{-1}\, \partial_y\, \partial_x (xc). \end{array} \right\}$$

for the metrics (4.32). Inspection shows that both classes are of type III,(or more special), so that we have proven the formal property "$p = 1$". We claim that "$p = 1$" is equivalent to the existence of plane wave surfaces: of course, for $p = 1$ the 2-surfaces $(v,u) = \mathrm{const}$ are flat. Conversely, every wave surface W can be described by $u = \mathrm{const}$, $v = f\,(x^\alpha)$, $\alpha \neq 3$, and its induced metric is

$$(5.14)\qquad G_2 = p^2\,|dz|^2 \quad ;$$

if W is plane, Cartesian coordinates can be introduced by means of a transformation (4.10) III, and this for every value of u ; which proves that $p = 1$ can be achieved.

Proceeding towards simpler and simpler subclasses, we state:

Theorem 5.3: The fields of type N and O are characterized as the only expansion-free radiation fields which possess plane and geodetic wave surfaces. They were obtained in (4.11), ($\Omega = 0$), and (4.32), ($\Omega \neq 0$), for respectively $b_{,AB} = 0$ and $(xc)_{,AB} = 0$; and could be further simplified to $b = 0$ and $c = 0$. We call them "plane-fronted".

Proof: The first statement will be subject of § 7. The

second statement can be gathered from (5.12) and (5.13), respectively: (5.5) holds good in (5.12) if and only if $b_{,Ax} = 0$, and together with $\Delta b = 0$ we arrive at $b_{,AB} = 0$; ~~and together with $\Delta b = 0$ we arrive at $b_{,AB} = 0$~~; in (5.13) the condition is "$\chi_a = 0$" , which together with $\Delta\, xc = 0$ amounts to $(xc)_{,AB} = 0$. These are the conditions (4.12) and (4.33) respectively which were recognized to be necessary and sufficient for resp. $b = 0$ and $c = 0$.

<u>Theorem 5.4:</u> The subclass (4.11) of expansion-free radiation fields is characterized by the <u>vanishing</u> of the <u>rotation</u> $|\Omega|$ of its ray congruence; that is the <u>rays</u> are <u>parallel</u>. This subclass is of type III, N, or O.

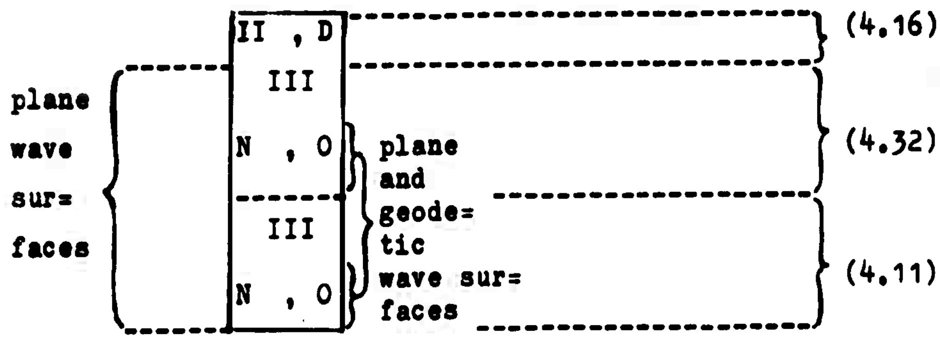
Proof: The proof was already given in (5.7) and (5.12). By "parallel rays" we mean that

$$\overset{\xi}{\nabla} l^a \sim l^a \tag{5.15}$$

holds for every vector ξ^b , so that the tangent vector l^a is parallelly transferred in every direction. (5.6) shows that within our class , $\tau_a = 0$ is characteristic for parallel rays.

We collect some of our results obtained up to this time in the following

<u>diagram of the expansion-free radiation fields:</u>



The careful reader has observed that our formulae (5.6) to (5.13) contained a great deal of geometric information, and that the coordinate systems introduced in § 4 are highly intrinsic; compare (4.10). As an example we mention:

Theorem 5.5 : For the metrics (4.16), $|\Omega|$ and J (in (5.10)) are scalar invariants.[19] In (4.32) , x is an invariant coordinate via $\sqrt{2}\,x = |\Omega|^{-1}$; y is invariant up to a u-dependant translation; the degree of arbitrariness of v depends on how far one can specify c (in general, f in (4.10) II is cut down to $f = a(u)\,x^2 + b(u)\,x + c(u)\,y + d(u)$); and u is unique up to an arbitrary scale change I.

The proof is almost obvious. For example, y is intrinsically distinguished by: $y_{,a}\,l^a = 0 = y_{,a}\,x'^a = y_{,a}\,v'^a$, $y_{,a}\,y'^a = 1$, in which equations only v is not invariant, resulting in an undeterminacy of $y_{,a}$ by an additive multiple of l_a .
If the field is conformally flat we have

---

19) Flat space-time is excluded.

$\mu \neq 0$, and again the ray congruence is intrinsically determined (by R_{ab}) . We have ignored the possible twofold ambignity of l_a for D-fields with $\mu = 0$.

The N-class of plane-fronted waves with parallel rays appearing in the bottom line of our diagrams is the simplest and best understood subclass. It has been subject of many papers by various authors, and is thoroughly treated in our article[18]. For the sake of completeness we quote some of its characteristic properties in

<u>Theorem 5.6:</u> Among all fields satisfying equations (4.8) the <u>plane fronted waves with parallel rays</u> (= pp waves) are characterized by one of the following properties: <u>1</u>) They are the <u>purely transverse</u> expansion-free radiation fields with <u>non-rotating</u> rays. <u>2</u>) They possess a covariantly constant null vector; (for vacuum fields , ($\mu = 0$), the property "null" is a consequence). <u>3</u>) They possess a covariantly <u>constant bivector</u> ; (this bivector is necessarily null). <u>4</u>) Their conform tensor is <u>complex recurrent</u> : put $C_{abcd} + i\,{}^{*}C_{abcd} = \Gamma_{abcd}$, then $\Gamma_{abcd;e} = \Gamma_{abcd}\, q_e$. <u>5</u>) Its <u>infinitesimal holonomy group</u> is 2-dimensional.[20]

Explanation: the metric fields is called "purely transverse" in a point x if there exists , in x , a null vector l_a such

---

20) This property was proved by J.N. Goldberg and R.P. Kerr, Aeron. Res. Lab., Ohio 1960.

that relative accelerations and relative rotations of inertial directions are orthogonal on the space-projection of l_a into the 3-space of every observer. The constant null bivector of the pp waves will be discussed below in connection with the MAXWELL field.

As was mentioned in 2) of the last theorem, pp waves admit a constant null vector, that is especially a group G_1 of isometries with lightlike orbits. The question arises whether this weaker property is already characteristic. The answer is negative, and gives further insight:

Theorem 5.7 :[21] The only solutions of (4.8) which admit a G_1 with lightlike orbits are 1) the pp-waves and 2) a class of II (D)-fields which are given by (4.16) together with

$$\Delta b^{2/3} = 0 \quad , \tag{5.16}$$

and the rest of (4.8).[22]

Theorem 5.8 : The metrics (4.16), (5.16) of the preceding theorem include the fields (4.20), which contain, as a simplest representative, the static vacuum field

$$G = x^{-1/2} | dz - v x^{-1/2} du |^2 + 2dv\, du - v^2 x^{-3/2} du^2 . \tag{5.17}$$

This last field possesses a second expansion-free ray congruence spanned by the gradient of $v x^{-1}$; it is therefore of type D.[23]

---

21) The corresponding theorem 3.3 in [6] is wrong.

22) The last line of (4.16) is here identically fulfilled!

23) It is the field R 3 in the classification of [18].

Proof of theorem 5.7: Consider a space-time V_4 which admits a null KILLING vector k_a:

(5.18) $$k^a k_a = 0 = k_{(a;b)} \quad .$$

k_a has vanishing propagation derivative: $\overset{k}{\nabla} k_a = 0$, hence we can use the decomposition formula (2.2), and obtain:

(5.19) $$k_{a;b} = 2\, i\, \omega\, \bar{t}_{[a}\, t_{b]} + 4\, \mathrm{Re}(\Omega\, t_{[a} k_{b]})$$

because of $k_{(a;b)} = 0$. But $\theta = 0 = \sigma$ implies $\omega = 0$ for $R_{ab}\, k^a k^b = 0$, (compare (2.15)), and ~~from~~ (5.19) simplifies to:

(5.20) $$k_{a;b} = \tau_{[a} k_{b]} \quad \text{with} \quad \tau^a k_a = 0 \; .$$

We therefore have $k_{[a;b}\, k_{c]} = 0$, so that there exists a scalar e^{τ}, say, with

(5.21) $$e^{\tau}\, k_a = u_{,a} \quad .$$

Insertion into (5.20) yields

(5.22) $$\tau_a = \tau_{,a} + \varkappa\, k_a \quad , \text{and}$$

(5.23) $$u_{;ab} = u_{,(a}\, \tau_{,b)} \quad \text{with} \quad \tau_{,a}\, u^{,a} = 0.$$

We have found that k_a is tangent to an expansion-free ray congruence, and that moreover $l_a \equiv u_{,a}$ satisfies (5.23).

A comparison with (5.6), (5.7) gives:

(5.24) $$\tau_{,a} \overset{!}{=} a_{a,3} \quad ,$$

so that (5.23) becomes equivalent with the integrability conditions

(5.25) $$\partial_3\, a_{[a,b]} = 0 \; .$$

Now again we have to distinguish between $= 0$ and $\neq 0$. In the

first case, (5.25) is equivalent with (4.12), so that we deal with the null subclass (pp waves). For the rest of the proof we assume $\Omega \neq 0$; we therefore have (compare (4.16), (4.14)):

(5.26) $$m_{A,3} = p^2 (b_{,1}{}^2 + b_{,2}{}^2)^{-1} b_{,A} , \quad m_3 = 1 ,$$

and $R_{ab} l^a m^b = 0$ amounts to:

(5.27) $$m_{4,33} = \frac{1}{2} p^{-2} (m_{1,3}{}^2 + m_{2,3}{}^2 - m_{1,31} - m_{2,32}) .$$

Now we find from (5.24) and (5.26):

(5.28) $$\left\{ \begin{array}{l} m_{a,3} = \partial_a [f(b) + g(u)] , \text{ with} \\ f' = p^2 (b_{,1}{}^2 + b_{,2}{}^2)^{-1} . \end{array} \right\}$$

Next we get from (5.25) : $0 = m_{4,33}$, and (5.27) yields:

(5.29) $$f_{,AA} = f_{,B} f_{,B}$$

with summation over repeated indices. Or more explicitly:

(5.30) $$f'' b_{,A} b_{,A} + f' b_{,AA} = f'^2 b_{,B} b_{,B} .$$

We now use one more field equation (taken from (4.16)):

(5.31) $$b_{,AA} = - \frac{1}{2} p^2 ,$$

and insert it in (5.30) to obtain:

(5.32) $$(f'' - \frac{3}{2} f'^2) b_{,A} b_{,A} = 0 ,$$

where use was made of (5.28) . Here the first factor must vanish because of $\Omega \neq 0$. Neglecting trivial constants of integration, we get

(5.33) $$f = - \frac{2}{3} \ln | b | .$$

With this result we enter into (5.28) to eliminate f:

(5.34) $$p^2 = -\frac{2}{3b} b_{,A} b_{,A} ,$$

and find that (5.31) becomes equivalent with (5.16).

In order to satisfy the last of equations (4.16) we use it in the shape of (4.15) , and apply (5.27) to eliminate the derivatives of $m_{A,3}$. We receive

(5,35) $$\Delta \ln p = \frac{1}{4} p^4 (b_{,A} b_{,A})^{-1} .$$

Astonishingly enough, this equation is identically satisfied by (5.34) with (5.16) : In order to see this the substitution $b^{2/3} = F$ is useful, for it turns (5.35) into

(5.36) $$\Delta \ln F_{,A} F_{,A} = 0 ;$$

which holds true in virtue of F being a potential function.

So far we are at the end of our proof. We have not yet taken account of the three field equations $R_{ab} m^a t^b = 0$, $2 R_{ab} m^a m^b = -\mu$,(which are messy), nor have we considered the two final conditions (5.25) :

(5.37) $$\partial_3 m_{[A,4]} = 0 .$$

The example of theorem 5.8 suggests, however, that no contradiction is to be expected, At least our class 2) of theorem 5.7 is not empty.

We are not willing to present the calculations leading to theorem 5.8. That (5.17) is a <u>static</u> metric can be seen by introducing $\bar{v} = v x^{-1}$ as a new coordinate:

(5.38) $$G = x^{-1/2}\,|dz|^2 + 2x\,d\bar{v}\,du \quad ;$$

here $\delta_3^a - \delta_4^a$ is a timelike, hypersurface-orthogonal KILLING vector.

The rest of this paragraph shall be devoted to MAXWELL's field Φ_{ab} which gives rise to $\mu \neq 0$ in the combined EINSTEIN-MAXWELL equations (4.1), (4.8). That is we consider electromag= netic waves which travel shoulder by shoulder with pure gravitational waves; (the latter represented by the relative acceleration field determined through the conform tensor). As was already pointed out in § 4, such waves have to obey the equations:

(5.39) $$\begin{cases} -2\,R_{ab} = \Phi_{ac}\,\bar{\Phi}_b{}^c = k_a\,k_b\ , \ k^a k_a = 0 \\ \Phi^{ab}{}_{;b} = 0 \end{cases} ,$$

whereby the first line fixes Φ_{ab} to be of the form (4.6):

(5.40) $$\Phi_{ab} = 2\,e^{i\varphi}\,k_{[a}\,t_{b]}\ ,\ \varphi \text{ real},\ k^a t_a = 0,\ t^a \bar{t}_a = 1;$$

here t_a can be chosen as one likes. Now the second line is a linear system of first order differential equations for φ (which need not have a solution). For $\omega = 0$, the general solution φ is obtained from a special one, φ_0 , by adding an arbitrary function of a phase u , where

(5.41) $$k_a = \varkappa\,u_{,a} \ .$$

(For a proof, assume φ_i , $i = 1,2$, to be two solutions of

(5.39) , (5.40). Their difference $\Psi \equiv \varphi_1 - \varphi_2$ satisfies $\phi^{ab}\Psi_{,b} = 0$, whence: $\Psi_{,b} = \alpha u_{,b}$, and necessarily $\alpha = \alpha(u)$.). We are therefore only forced to find special solutions. Let us try to find such solutions for the expansion-free radiation fields!

Consider our metrics (4.11), (4.16), (4.32). In all of them μ was left arbitrary except for

(5.42) $\dot{\mu} = 0$; $(\mu = \varkappa^2)$.

It now turns out that μ has to satisfy a non-linear third order differential equation in order to allow for a solution (5.39). This equation admits as solution an arbitrary function of u , and in this special case we can easily solve our problem. If, however, μ depends on the spatial coordinates x^A , we can only present special solutions. We claim:

Theorem 5.9: For $\overset{t}{\nabla}\mu = 0$, the general MAXWELL bivector in an expansion-free radiation field (4.11), (4.16), (4.32) is given by

(5.43) $$\phi_{ab} = 2\, e^{i\varphi(u)}\partial_{[a} f(u)\partial_{b]} z \quad \text{with } f'^2 = \mu(u) , \quad f, \varphi \text{ real} .$$

(Note the special case of a pp wave with $\mu_{,a} = 0 = \varphi_{,a}$ in which ϕ_{ab} is constant!) .

For $\overset{t}{\nabla}\mu \neq 0$, ϕ_{ab} must necessarily have the shape $(\mu = \varkappa^2)$:

(5.44) $$\phi_{ab} = \sqrt{2}\, e^{\lambda + i\varphi} (k_{[a,b]} + i\, k^{*}_{[a,b]}),$$

$$k_{[a,b]} = u_{,[a}\, \varkappa_{,b]}\ , \quad e^{\lambda} = \varkappa (\varkappa^{,A} \varkappa_{,A})^{-1/2}$$

and (5.39) become equivalent with

(5.45) $$\begin{cases} \varphi_{,A} = -\, \varepsilon_{AB} \{ \varkappa_{,B}\, \Delta \varkappa (\varkappa_{,1}{}^{2} + \varkappa_{,2}{}^{2})^{-1} + \lambda_{,B} \} , \\ \varepsilon_{AB} = \begin{cases} +1 & \text{for } (AB) = (12) \\ -1 & \text{for } (AB) = (21) \\ 0 & \text{" otherwise} \end{cases} , \end{cases}$$

whose integrability condition $0 = \varphi_{,[AB]}$ is of third order in $\varkappa$. (5.45) is solved by $\varphi = 0$ if and only if

(5.46) $$e^{\lambda} = f'(\varkappa) \quad \text{with } \Delta f = 0;$$

for $p = 1$ a special solution of this kind is given by

(5.47) $$\varkappa = | a(u) \quad z + b(u) |^{-1} \quad , \text{ a,b complex} .$$

Proof: (5.43) clearly satisfies the first line of (5.39) . The second line of (5.39) is equivalent with $\phi_{[ab,c]} = 0$ because of $\overset{*}{\phi}_{ab} = -\, i\, \phi_{ab}$, and is therefore obviously fulfilled. (5.43) is the general solution (in the case $\overset{*}{\nabla} \mu = 0$) because φ depends arbitrarily on u (compare our remark above).

We come to the case $\overset{*}{\nabla} \mu \neq 0$. Here again it is not difficult to see that (5.44) is of the form (5.40). In order to satisfy the second line of (5.39) we use the formula

(5.48) $$\eta_{abcd} = 4!\, \xi_{[a} \eta_{b}\, l_{c}\, m_{d]} \ ,$$

with $\sqrt{2}\,t_a = \xi_a + i\,\eta_a$, $l_a = u_{,a}$, $l^a\,m_a = 1$, to find

$$(5.49)\quad \left\{\begin{array}{l} k_{[a,b]} + i\,\overset{*}{k}_{[a,b]} = u_{,[a}\, s_{b]} \text{ with} \\ s_a = \varkappa_{,a} - i\,\delta_a^A\, \varepsilon_{AB}\, \varkappa_{,B} \end{array}\right\};$$

(summation over dummy indices). We thus get from (5.44) ,

(remember: $l^a = \delta_3^a$; and $F^{ab}{}_{;b} = g^{-1/2}(g^{1/2}F^{ab})_{,b}$

for every bivector F^{ab}):

$$\sqrt{2}\,\phi^{ab}{}_{;b} = e^{\lambda+i\varphi}\left\{ l^a\left[\overset{s}{\nabla}(\lambda + i\,\varphi) + \partial^B s_B\right] - s^a\,\overset{1}{\nabla}\, i\varphi\right\},$$

and (5.39) yields:

$$(5.50)\qquad \dot{\varphi} = 0\,,\ \overset{s}{\nabla}\varphi = i\,(\overset{s}{\nabla}\lambda + \partial^B s_B)\,.$$

This last equation reads in matrix form:

$$(5.51)\quad \begin{pmatrix} \varkappa_{,1} & \varkappa_{,2} \\ -\varkappa_{,2} & \varkappa_{,1} \end{pmatrix}\begin{pmatrix} \varphi_{,1} \\ \varphi_{,2} \end{pmatrix} = \begin{pmatrix} \varkappa_{,2}\,\lambda_{,1} - \varkappa_{,1}\,\lambda_{,2} \\ \varkappa_{,1}\,\lambda_{,1} + \varkappa_{,2}\,\lambda_{,2} + \Delta\varkappa \end{pmatrix},$$

which is easily solved, with (5.45) as the result.

The rest of theorem 5.9 is understood as follows: if $\varphi = 0$ is to be a solution of (5.51) , the first line tells us that λ must be a function of $\varkappa$; we write $e^{\lambda} = f'(\varkappa)$. But then the second line, multiplied by e^{λ} , gives $\Delta f = 0$. Conversely, starting with a suitable potential function f, one can try to solve $e^{\lambda} = f'(\varkappa)$ (with e^{λ} given in (5.44) , and $\varkappa = F(f)$). In this way one easily finds the solution (5.47) , with $|a|\,f = \ln\varkappa = -\ln|az + b|$.

§ 6 Plane-fronted waves in MAXWELL's theory.

In this paragraph we want to study those electromagnetic wave fields in flat space-time which one would intuitively call plane-fronted: namely the waves emitted by an idealized searchlight. Contrary to our previous manipulations we will this time use MINKOWSKIAN coordinates. The reason for presenting an intuitive derivation is twofold: 1) to penetrate our formal operations of above with physical understanding, and 2) to win a reasonable formal definition for what one should call "plane-fronted" in a curved V_4 .

Consider an electromagnetic point source in flat space which is located in the focus of a good lens. The waves which we find behind the lens shall be called "plane-fronted" even if the lens is allowed to move at constant distance from the source. We are going to derive the general MAXWELL field of this kind: Consider a (plane) wave front which at time t is just about to leave the lens. Its history is described (in MINKOWSKIAN coordinates, $c = 1$), by:

(6.1) $\qquad S(x^{\alpha}; u) - t = 0 \, , \quad S^{,\alpha} S_{,\alpha} = 1, \; (\alpha = 1,2,3) \, ,$

where S is the eikonal function which depends linearly on x^{α} . The parameter u, a co-moving (null) coordinate, is designed

to number successive wave surfaces; necessarily $S_{,u} \neq 0$ for $x^\alpha = 0$.

Taking for u the retarded time, (6.1) reads more explicitly:

(6.2) $\quad x^\alpha s_\alpha(u) + u - t = 0\ ,\ s^\alpha s_\alpha = 1.$

Solving with respect to u we get $u = u(x^a)$,

with

(6.3) $\quad (1 + x^\alpha s_\alpha') \; u_{,a} = \delta^4_a - \delta^\beta_a s_\beta\ ,$

whence:

$$(6.4)\quad \left\{\begin{array}{l} u_{,a}u^{,a} = 0\ ,\ u_{,ab} = u_{,(a}\tau_{,b)} + \beta u_{,a}u_{,b} \\ \tau = -2\ln(1 + x^\beta s_\beta')\ ,\ \tau_{,a}u^{,a} = 0\ , \\ \tau^{,a}{}_{a} + \frac{1}{2}\tau^{,a}\tau_{,a} = 0. \end{array}\right\}$$

Note the special case of a resting lens (<u>parallel rays</u>):

$$s_\alpha = \text{const}\ ,\ \Rightarrow\ \tau_{,b} = 0 = \beta = u_{,ab}\ .$$

Formula (6.4) shows that the vector $u_{,a}$ describes an expansion-free ray congruence for which one can always find an accompanying null MAXWELL field (compare our remarks in connection with (4.6).) The latter field is unique ($\omega = 0$) up to a u-dependent "amplitude factor" e^λ, and a u-dependent "duality rotati..." of angle φ (in our notation of (5.44)). We are going to present this field explicitly: The ansatz

$$(6.5)\quad \left\{\begin{array}{l} \varkappa u_{,i} = k_i\ ,\ F_{ij} = 2k_{[i,j]}\ , \\ \text{with}\quad \dot\varkappa = \varkappa_{,i}u^{,i} = 0\ , \end{array}\right\}$$

for the electromagnetic null field F_{ij} satisfies MAXWELL's

equations $F_{[ij,k]} = 0 = F^{ij}{}_{,j}$ if and only if

$$(6.6) \qquad (e^{\tau/2} \varkappa)^{,j}{}_{j} = 0 \;;$$

and one can easily make sure that it allows for the generality (ambignity) described above; ((6.6) admits two solutions whose gradients are spacelike and orthogonal on each other). We have shown:

<u>Theorem 6.1</u> Equs. (6.2), (6.5), and (6.6) describe the general plane-fronted MAXWELL wave; amplitude and polarization are determined by one (real) function $\varkappa$ which satisfies a generalized homogeneous wave equation, and which is constant along the rays.

A <u>comparison</u> of (6.4) with (5.13) shows:

$\tau \;\leftrightarrow -2\ln x \;, \Rightarrow e^{\tau/2} \leftrightarrow x^{-1}$ for $\tau_{,b} \neq 0$. In order to compare (6.6) with (4.32), we have to write both formulae covariantly: In (6.6), of course, commas can be replaced by semicolons. But for the general metric (4.32) one gets

$$(6.7) \qquad x^{-k} \nabla^{a}_{a} x^{k} a = x^{-(k+1)} \Delta x^{k+1} a \;, \text{ if } a_{,v} = 0,$$

where use has been made of

$$(6.8) \qquad \begin{pmatrix} p^2 & & & m_1 \\ & p^2 & & m_2 \\ & & 0 & 1 \\ m_1 & m_2 & 1 & 2m_4 \end{pmatrix}^{-1} = \begin{pmatrix} p^{-2} & & -p^{-2}m_1 & \\ & p^{-2} & -p^{-2}m_2 & \\ -p^{-2}m_1 & -p^{-2}m_2 & -H & 1 \\ & & 1 & \end{pmatrix}$$

for $p = 1$, $m_2 = 0$. Now put $c = 0 = \mu$ in (4.32) to obtain

the general plane-fronted gravitational wave with rotating rays. It is described by one function $\underline{A}$ satisfying the same equations as the scalar $\underline{e^{-\tau/2}\varkappa}$ which describes the rotating plane-fronted electromagnetic wave field! An even simpler correspondence prevails between (6.5), (6.6) and (4.11), that is between the subclasses with parallel rays. This fact reveals a <u>one-to-one correspondence</u> between plane-fronted electromagnetic waves and plane-fronted gravitational waves (if theorem 5.3 is admitted). In this correspondence, however, one must count "modulo" constant electromagnetic fields, as the corresponding gravitational fields are flat; in agreement with the fact that (constant) accelerations have no invariant meaning in a general space-time, but only relative accelerations of neighbouring objects.

§ 7. The plane-fronted radiation fields

We are going to collect a number of properties of those MAXWELL fields which we called "plane-fronted" in the preceding paragraph into the

<u>Definition 7.1</u> A V_4 of signature $+2$ satisfying the fields equations (4.8) is called a <u>plane-fronted</u> wave if it has the following properties: 1) V_4 contains a ray congruence (= geodetic, distortion-free null congruence). 2) V_4 admits

a 2-parametric set of simple coverings by plane and geodetic wave surfaces.

The coverings described under 2) were given in (6.1) by $u = \text{const}$, $S - x^{\alpha} v_{\alpha} = \text{const}$, where v_{α} corresponds to an arbitrary velocity of the (rigid) system of observers; (S is supposed to transform like a scalar under frame changes, whereas t of course does not.- Only if s_{α} is linearly independent of the v_{α} do we get different coverings[24], hence only a 2-parametric set). In this paragraph we want to construct the general plane-fronted wave; (as we do not exclude flat space the congruences (6.4) will be contained in our solutions). The result was anticipated in theorem 5.3: the plane-fronted waves are the expansion-free radiation fields of type N or O .

Proof: Consider a plane-fronted wave in the sense of the above definition. From lemma 2.1 and the assumed existence of wave surfaces we get that the ray congruence is also normal. Consequently, the fundamental form can be brought to the form (3.9) or (3.12), that is to the form

(7.1) $$G = p^2 \, |dz|^2 + 2 \, m_a dx^a \, du \, , \; p m_3 \neq 0.$$

(We prefer to dispense with the gauge "$m_3 = 1$" in the following

24) Observers travelling in the direction of the wave normal agree in their definitions of "wave surfaces at fixed time".

derivation, as we intend to specialize the shape of the eikonal function). For these metrics the assumed set of coverings must be expressible in the form (compare the definition of wave surfaces in § 2):

$$(7.2)\qquad \left\{ \begin{array}{l} x^4 \equiv u = \text{const} \\ S\,(x^a\ ; a^K) = \text{const} \quad , \ (K = 1,2) \end{array} \right\}, \text{ with}$$

$$(7.3)\qquad \text{rank}\left(\frac{\partial^2 S}{\partial x^\alpha \partial a^K}\ ,\ \frac{\partial S}{\partial x^\alpha}\right) = 3,\quad S_{,3} \neq 0 \quad \text{for } a^K = 0,$$

$(\alpha = 1,2,3)$. (7.3) expresses the linear indapendence of the normals to the original surface $\left[(u,\ S\ ) = \text{const} \ \text{ for } \ a^K = 0\right]$, and to two neighbouring ones $\left[\delta a^K \neq 0\right]$; that is the fact that the parameters a^K are essential. By a suitable (in general not affine) gauge of v we can effect:

$$(7.4)\qquad S = v + O(a^K)\ .$$

The geodesies of (7.1) are the solutions of the LAGRANGEan system belonging to

$$(7.5)\qquad L = p^2 |\dot z|^2 + 2\, m_a \dot x^a\, \dot u\ ,\ (\dot{} \equiv \frac{d}{d\tau}\ ).$$

A geodesic of the hypersurface $u = \text{const}$ can only be geodesic of (7.1) if $(p^2)_{,v}|\dot z|^2 = 0$, that is if p is independent of v . The surfaces for $a^K = 0$ shall be plane, $\Rightarrow$ p can be made 1 by means of a gauge (4.10) III. With this the geodesics of the surfaces (7.2) are described by

$$(7.6)\qquad \dot u = 0 = \ddot z = \dot S\ .$$

They are assumed to solve the LAGRANGE equations belonging to (7.5) with $p = 1$; this demand amounts to the one equation

$$(7.7) \quad \left\{ \begin{aligned} 0 &= \frac{d}{d\tau}\,(m_A \dot{x}^A + m_3 \dot{v}) \\ &= m_{A,B}\,\dot{x}^A \dot{x}^B + 2\, m_{(A,3)}\,\dot{x}^A \dot{v} + m_3 \ddot{v} + m_{3,3} \dot{v}^2 \\ &\text{together with (7.6) and (7.4), identically in } a^K. \end{aligned} \right\}$$

We are going to evaluate (7.7).

1. $a^K = 0$: (7.7) implies $m_{(A,B)} = 0$, $\Rightarrow$

$$(7.8) \quad \left\{ \begin{aligned} m_1 &= -y\, g(v,u) + h_1(v,u) \\ m_2 &= x\, g(v,u) + h_2(v,u) \end{aligned} \right\},$$

and the term $m_{A,B}\,\dot{x}^A \dot{x}^B$ in (7.7) vanishes.

2. First order in a^K : We expand S :

$$(7.9) \quad S = v + a^K k_K (x^a) + O(a^K a^L),$$

and obtain from (7.7):

$$(7.10) \quad 0 = a^K x^L x^M (n_L k_{K,M} + m_3 k_{K,LM}) + O(a^K a^L),$$

where $n_A \equiv 2\, m_{(A,3)}$. Using this equation for $L = M$ one finds

$$(7.11) \quad k_{K,L} = l_{KL}(v,u)\, p_L (x^a),$$

whence $p_L = p_{,L}$ without restriction of generality (put $K = 1$). The integrability condition $0 = 2\, k_{K,[LM]} = (l_{KL} - l_{KM})\, p_{,LM}$ implies either $p_{,KL} = 0$ for $K \neq L$, or $l_{KL} = l_K$. The latter case violates the rank condition

(7.3), ergo

$$k_{K,12} = 0 \,. \tag{7.12}$$

Now (7.10) for $(L,M) = (1,2)$ implies

$$k_{K,(1}n_{2)} = 0 \tag{7.13}$$

with the only solution $n_L = 0$ because of $\det(k_{K,L}) \neq 0$ (on account of (7.3)). The integrability condition $m_{3,[AB]} = 0$ of $n_A = 0$ leads to

$$\left\{ \begin{array}{ll} g(v,u) = g(u) \\ m_{3,A} = -h_{A,3}(v,u) \end{array} \right\} . \tag{7.14}$$

3. Last step: (7.7) has now been reduced to

$$\left\{ \begin{array}{l} 0 = \partial_3 \ln m_3 - \frac{d}{d\tau}(\dot{v}^{-1}) \\ \text{with (7.6)}\ , \text{identically in } a^K \end{array} \right\} . \tag{7.15}$$

By eliminating $\dot{v}$ from $0 = \dot{S} = S_{,A}\dot{x}^A + S_{,3}\dot{v}$ one gets

$$-\frac{d}{d\tau}(\dot{v}^{-1}) = (S_{,A}\dot{x}^A)^{-2}\dot{x}^B\dot{x}^C\{S_{,B}(S_{,3C} - S_{,33}S_{,C}S_{,3}^{-1}) - S_{,3}(S_{,BC} - S_{,B3}S_{,C}S_{,3}^{-1})\} \tag{7.16}$$

which has to hold for arbitrary $\dot{x}^A$. The $\dot{x}^A$ can therefore be "divided out", and insertion into (7.15) yields

$$\partial_3\{\ln m_3 + \ln(S_{,B}S_{,C}S_{,3}^{-1})\} = S_{,3}S_{,BC}(S_{,B}S_{,C})^{-1} ,$$

or:

$$S_{,BC} = \partial_3 S_{,B}S_{,C}S_{,3}^{-1} + S_{,B}S_{,C}S_{,3}^{-1}\partial_3 \ln m_3 \,. \tag{7.17}$$

Here we have the integrability conditions

$$(7.18) \qquad 0 \overset{!}{=} S_{,A[BC]} = S_{,3}^{\,-1}\, S_{,A}\, S_{,[B}\partial_{C]}\partial_3 \ln m_3 + O(a^K a^L a^M)$$

which must hold identically in a^K, so that $\partial_C \partial_3 \ln m_3 = 0$, $\Rightarrow m_{3,C} = m_3 F_{,C}(x^A, u)$. From (7.14) we obtain

$$(7.19) \qquad -h_{A,3}(v,u) = F_{,A}(x^C, u)\left\{- h_{B,3}(v,u)\, x^B + h(v,u)\right\}.$$

which implies

$$(7.20) \qquad h_{A,3} = j_A(u)\, l_{,3}(v,u), \quad h = j_3(u)\, l_{,3}(v,u),$$

as can be seen by logarithmic differentiation of (7.19) with respect to x^B. Now we collect from (7.8), (7.14), (7.20):

$$(7.21) \qquad \left\{\begin{array}{l} m_1 = -y\, g(u) + j_1(u)\, l(v,u) + n_1(u) \\ m_2 = x\, g(u) + j_2(u)\, l(v,u) + n_2(u) \\ m_3 = -\left[j_A(u)\, x^A - j_3(u)\right] l_{,3}(v,u) \end{array}\right\}.$$

Instead of v, $l(v,u)$ can be introduced as the new variable $\bar{v}$; ($l_{,3} \neq 0$ holds according to (7.1)) . As none of the earlier assumptions is violated by this gauge, (7.21) is now valid with v instead of $l(v,u)$. The plane and geodetic coverings are then given by (7.2) with

$$(7.22) \qquad S = v + a^K x^K ,$$

they satisfy (7.17) identically, and all conditions of plane-frontedness are taken care of.

4. In order to show that (7.1) with $p = 1$ and (7.21) are the announced $N(0)$-fields, we simplify m_a by means of the gauges (4.10). In this process we make use of the fact

that (7.21) are form-invariant under the restricted class, for which f in II is linear in x, y, and $|F'| = 1$ in III; (because all presuppositions are form-invariant under this class). Two cases have to be distinguished, namely $\mathcal{R} = 0$ and $\mathcal{R} \neq 0$ as above.

a) $j_A = 0$. In this case $d\bar{v} = j_3(u)dv$ leads to $m_{\bar{3}} = 1$, and g and n_A can be gauged away with III. We therefore arrive at (4.11) with $b = 0$.

b) $j_A j_A \equiv k^2 > 0$. The succeeding gauges

1. $j_A(u)x^A - j_3(u) = k\bar{x}$, $kv = \bar{v}$, $(\Rightarrow \bar{j}_1 = 1, \bar{j}_2 = 0)$
2. $v - y\,g(u) = \bar{v}$ $\qquad (\Rightarrow \bar{g} = 0)$
3. $v + n_1(u) = \bar{v}$ $\qquad (\Rightarrow \bar{n}_1 = 0)$
4. $y + \int n_2(u)\,du = \bar{y}$ $\qquad (\Rightarrow \bar{n}_2 = 0)$

transform (7.21) into

$$G = |dz|^2 + 2\,(v\,dx - x\,dv + \tfrac{1}{2} H\,du)\,du\,. \tag{7.23}$$

From here the canonical form (4.32) with $c = 0$ is reached by the transformation $-xv = \bar{v}$, (which destroys the form (7.22) of the eikonal function). This remark finishes the proof.

CHAPTER IV. HYDRODYNAMICS

§ 1. Basic Concepts. Decomposition of the velocity-gradient

In the theory of relativity the history (motion) of a continuous medium is described by a congruence of timelike curves, the worldlines of the fluid elements. With respect to an arbitrary local coordinate-system let

(1.1) $$x^a = x^a(y^\alpha, s)$$

be a parameter-representation of the congruence; here (y^α) serves to identify a curve, and s measures proper time along it. Then

(1.2) $$u^a \equiv \frac{\partial x^a}{\partial s} \equiv \dot{x}^a \quad (u_a u^a = -1)$$

is the (four-) velocity field .

If

(1.3) $$\delta \equiv \delta y^\alpha \frac{\partial}{\partial y^\alpha}$$

denotes the variation transverse to the streamlines and a dot indicates covariant differentiation with respect to s along them we have

(1.4) $$(\delta x^a)^{\cdot} = u^a{}_{;b}\,\delta x^b ,$$

i.e. a connection-vector δx^a of two neighbouring curves is Lie-transferred.

Introducing the projection tensor

(1.5) $$h^a_b \equiv \delta^a_b + u^a u_b$$

we may form

(1.6) $\delta_{\perp} x^a = h^a_b \delta x^b$

which may properly be called the relative radius vector of the respective particles. The Fermi-derivative

(1.7) $v^a = h^a_b (\delta_{\perp} x^b)^{\cdot}$

of it is, by definition, the velocity of the particle $(y^\alpha + \delta y^\alpha)$ relative to (y^α).

The (absolute) accelerations of the particles are given by

(1.8) $\dot{u}^a = u^a{}_{;b} u^b \quad . \quad (u_a \dot{u}^a = 0)$

The five preceding equations allow to calculate the relative velocity field describing the flow near one particle; one obtains

(1.9) $v^a = u^a{}_{;b} \delta_{\perp} x^b .$

As in ordinary hydrodynamics we decompose the tensor $u^a{}_{;c} h^c_b$ effecting the transformation $\delta_{\perp} x^a \rightarrow v^a$ into irreducible parts:

(1.10) $u_{a;c} h^c_b = \omega_{ab} + \sigma_{ab} + \frac{1}{3} \Theta h_{ab})$

with

(1.11) $\omega_{(ab)} = \sigma_{[ab]} = 0 \ , \ \sigma^a{}_a = 0, \ \omega_{ab} u^b = \sigma_{ab} u^b = 0$

Accordingly the infinitesimal transformation (1.9) is a superposition of a rotation $\omega^a{}_b \delta_{\perp} x^b$, a volume-preserving deformation $\sigma^a{}_b \delta_{\perp} x^b$, and a dilatation $\frac{1}{3} \Theta \delta_{\perp} x^a$.

The distances $\delta l \equiv (g_{ab}\delta_{\perp}x^{a}\delta_{\perp}x^{b})^{\frac{1}{2}}$ of adjacent particles change, in consequence of eqs. (1.7,9,10,11), according to

$$(1.12) \qquad \frac{(\delta l)^{\cdot}}{\delta l} = \frac{1}{3}\Theta + \sigma_{ab}e^{a}e^{b} \quad \left(e^{a} \equiv \frac{\delta_{\perp}x^{a}}{\delta l},\ e_{a}e^{a} = 1\right),$$

and the directions pointing to the surrounding particles change as follows:

$$(1.13) \qquad h^{a}_{b}\dot{e}^{b} = (\omega^{a}_{b} + \sigma^{a}_{b} - \sigma_{cd}e^{c}e^{d}\delta^{a}_{b})e^{b}.$$

Eq. (1.12) implies

$$(1.14) \qquad \Theta = 3\overline{\frac{(\delta l)^{\cdot}}{\delta l}} = \frac{(\delta V)^{\cdot}}{\delta V}$$

if a bar denotes averaging and δV is the volume of a small piece of the medium.

According to (1.13), $h^{a}_{b}\dot{e}^{b} = \omega^{a}_{b}e^{b}$ holds precisely if e^{a} coincides with a main direction of shear $(\sigma^{a}_{b}e^{b} \sim e^{a})$. Consequently, that material orthogonal three-leg which coincides at a certain time s with the eigen three-leg of the shear velocity tensor σ^{a}_{b} undergoes, during $(s, s+ds)$, the infinitesimal rotation determined by ω^{a}_{b}.

From equs. (1.10,11) we obtain the explicite expressions

$$(1.15) \qquad \omega_{ab} = u_{[a;b]} + \dot{u}_{[a}u_{b]},$$

$$(1.16) \qquad \sigma_{ab} = u_{(a;b)} + \dot{u}_{(a}u_{b)} - \frac{1}{3}\Theta h_{ab},$$

$$(1.17) \qquad \Theta = u^{a}{}_{;a}$$

for the vortex-tensor, the shear tensor, and the expansion scalar, respectively. ($\sigma_{ab} + \frac{1}{3}\Theta h_{ab}$ is the rate-of-strain tensor.)

In place of ω_{ab} the vortex vector (or angular velocity)

(1.18) $$\omega^a = \frac{1}{2}\eta^{abcd} u_b \omega_{cd} = \frac{1}{2}\eta^{abcd} u_b u_{c,d}$$

can be used; in the local rest frame defined by u^a ω^a is the (spatial) dual of ω_{ab} :

(1.19) $$\omega_{ab} = \eta_{abcd}\,\omega^c u^d \quad (\Rightarrow \omega_{ab}\omega^b = 0)$$

The relation in brackets shows ω^a to determine the axis of rotation which is orthogonal to the plane defined in the tangent space by the simple, spacelike bivector ω_{ab} .

In general a timelike congruence has nine (independent) first order invariants: The six "spatial" components of $\dot{u}^a$ and ω^a with respect to the eigen-tetrad of $\sigma^a{}_b$, two independent eigenvalues of $\sigma^a{}_b$ (or, equivalently, $\sigma^a{}_b\sigma^b{}_a$ and $\sigma^a{}_b\sigma^b{}_c\sigma^c{}_a$), and Θ . These quantities determine (if the metric is given) the congruence uniquely up to homogeneous LORENTZ-transformations within the infinitesimal neighbourhood of first order of an event x^a as can be seen from the relation

(1.20) $$u_{a;b} = \omega_{ab} + \sigma_{ab} + \frac{1}{3}\Theta h_{ab} - \dot{u}_a u_b$$

which follows by adding (1.15) and (1.16).

It is useful to introduce the absolute magnitudes

(1.21) $\quad \dot{u} \equiv (\dot{u}_a \dot{u}^a)^{1/2}, \; \omega \equiv (\omega_a \omega^a)^{1/2} = (\frac{1}{2}\omega_{ab}\omega^{ab})^{1/2},$

$$\sigma \equiv (\tfrac{1}{2}\sigma_{ab}\sigma^{ab})^{1/2}$$

of acceleration, rotation and shear which are non-negative and vanish only if the corresponding tensors do.

It should be stressed that the preceding definition of rotation rests upon the convention to use Fermi-propagated axes as the standard of rest. The dynamical meaning of this rotation has been illustrated by idealized experiments with test bodies (compare, e.g., PIRANI 1956), SYNGE 1960) and is implicitly contained in every theorem following from the field equations in which ω^a occurs; examples are given below.

2. Some special kinematical relations

From the works of SYNGE 1937, 1960, LICHNEROWICZ 1955, RAYNER 1959 and others the differential-geometrical meaning of special flows such as, e.g., irrotational, incompressible, rigid motions and some theorems about them are known. This section is devoted to some additional statements of this type which are independent of field equations.

1. According to LICHNEROVICZ, an irrotational $(\omega = 0)$ flow is volume-preserving $(\theta = 0)$ if and only if the hypersurfaces orthogonal to the streamlines (which exist according to $\omega = 0$,

see eq. (1.18)) are minimal. A similar statement which is useful for a geometrical theory of static space-times[1] is this:

<u>theorem</u>: An irrotational flow ($\Leftrightarrow$ normal congruence, SYNGE 1937) is rigid if and only if the hypersurfaces orthogonal to the stream-lines are totally geodetic.

<u>2.</u> It has been shown by SALZMAN and TAUB 1957 that the streamlines of a flow are the trajectories of a one-dimensional group of isometries if and only if the flow is rigid and admits an acceleration-potential U ($\dot{u}_a = U_{,a}$). A generalisation is:

<u>theorem:</u> The streamlines of a flow are the trajectories of a one dimensional (local) group of conformal mappings of space-time onto itself if and only if the flow is shear-free ($\sigma = 0$) and $\dot{u}_a$ admits a representation such as

(2.1) $$\dot{u}_a = \frac{1}{3}\Theta u_a + U_{,a} \; ; \; (\Rightarrow \dot{U} = \frac{1}{3}\Theta)$$

if these conditions are satisfied, $\xi^a = e^U u^a$ generates the corresponding group.

A simple example of such a flow is given by the motion of the substratum of a FRIEDMANN-LEMAITRE universe. In fact, the theorem

---

1) This will be presented in the second chapter by Ehlers and Kundt of the book on the theory of gravitation edited by L. Witten, to appear in 1961 (Wiley)

enables one to simplify the derivation and kinematical characterization of these universes. It should be noted that, by (2.1), a geodesic "conformal" flow (with $\Theta \neq 0$) is always irrotational.

3. It is well-known (e.g. EISENHART 1926) that the scalar product of the unit tangent vector of a geodesic with the generating vector of an isometry-group is constant along the geodesic. It is easily seen that this constancy also holds for a null-geodesic l and a conformal group if the tangent vector k^a of the former is chosen such that $k^a{}_{;b}\, k^b = 0$. Using the above notation we have

(2.2) $$e^U k_a u^a = \text{const. along } l\,.$$

Now, if the conformal congruence is interpreted as the set of worldlines of light-emitters and receivers, and the null geodesic as a light ray, (2.2) states

(2.3) $$e^u . \nu = \text{const. along } l,$$

where ν is the frequency "with respect to u^a".
This simple remark contains the theory of the frequency-shift in stationary and static gravitational fields and in the FRIEDMANN-LEMAITRE - universes. In the former we have $\Theta = 0$, $\dot{u} \neq 0$, and U is (by definition, but strongly suggested by (2.1) and further properties of U, see, e.g., eq. (4.3)) the scalar gravitational potential, and in the latter we have $\Theta \neq 0$, $\dot{u} = 0$, and e^U is (cf. (2.1)) proportional to the radius of curvature of the space.

3. Differential identities for the kinematical quantities

The derivatives of the quantities introduced in the first section satisfy, in consequence of the RICCI-identity for u^a, certain identities into which the curvature tensors enter. We shall now present them and give some applications.

We start by calculating the relative acceleration of neighbouring particles, defined-analogously to (1.7) - by

$$b^a = h^a_b \dot{v}^b. \tag{3.1}$$

Using (1.9) and the RICCI-identity we obtain

$$b^a = (R^a{}_{bcd} u^b u^d + h^a_b \dot{u}^b{}_{;c} + \dot{u}^a \dot{u}_c)\delta_\perp x^c \tag{3.2}$$

which generalizes the well known formula for the deviation of geodesics.

Next, we write down the RICCI-identity for u^a, expressing the first derivatives of u_a by means of (1.20). We obtain

$$\begin{aligned} \frac{1}{2} R_{abcd} u^d = {} & \omega_{c[a;b]} + \sigma_{c[a;b]} + \frac{1}{3} h_{c[a} \Theta_{,b]} - \dot{u}_{c;[b} u_{a]} \\ & + \frac{1}{3}\Theta(u_c \omega_{ab} - u_c \dot{u}_{[a} u_{b]} - u_{[a} \omega_{b]c} + \sigma_{c[b} u_{a]} + \frac{1}{3}\Theta g_{c[b} u_{a]}) \\ & - \dot{u}_c (\omega_{ab} - \dot{u}_{[a} u_{b]}). \end{aligned} \tag{1.3}$$

By contraction we get, after some calculation in which the relations of section 1 are used:

(1.4) $$R_{ab}u^a u^b = \dot{\theta} + \frac{1}{3}\theta^2 - \dot{u}^a{}_{;a} + 2(\sigma^2 - \omega^2),$$

$(-\sigma^{bc}{}_{;c}$

(1.5) $$h^{ab}R_{bc}u^c = h^a_b(\omega^{bc}{}_{;c} + \frac{2}{3}\theta^{,b}) + (\omega^a{}_b + \sigma^a{}_b)\dot{u}^b .$$

Finally we transvect (1.3) with u^b, symmetrize with respect to a and c and transvect with $h^a_e h^c_f$ in order to simplify the resulting expression:

(1.6) $$R_{abcd}u^b u^d = \omega_a \omega_c - \omega^2 h_{ac} + \sigma_{ab}\sigma^b{}_c + \frac{\ddot{l}}{l} h_{ac}$$
$$- \dot{u}_f \dot{u}_c + h_{ab}h_{cd}(l^{-2}(l^2\sigma^{bd})^{\cdot} - \dot{u}^{b;d})$$

Here l is an auxiliary "scale factor" defined (up to a scalar factor which is constant along the streamlines) by

(1.7) $$\frac{\dot{l}}{l} = \frac{1}{3}\theta , \quad \Rightarrow \frac{1}{3}(\dot{\theta} + \frac{1}{3}\theta^2) = \frac{\ddot{l}}{l} .$$

By antisymmetrizing $R_{abcd}u^b u^d$, calculated with (1.3), with respect to a,c one obtains a propagation-law for the vortex-vector, namely

(1.8) $$h^a_b(l^2\omega^b)^{\cdot} = \sigma^a{}_b l^2\omega^b + \frac{1}{2}{}^2\eta^{abcd}u_b\dot{u}_{c;d} .$$

A further identity satisfied by ω^a is

(1.9) $$\omega^a{}_{;a} = 2\dot{u}_a\omega^a ;$$

i.e. the vortex vector is not solenoidal as in ordinary hydrodynamics.

We now proceed to describe some applications of these relations.

At first we note two theorems which complement the vortex-theory developed by SYNGE 1937, LICHNEROVICZ 1955, and others.

<u>Theorem 3.1.</u> The vortex-lines (i.e. the curves with direction ω^a) and the streamlines generate surfaces if and only if the condition

(3.10) $$u^{[a}\omega^{b}\eta^{c]def}u_d\dot{u}_{e;f} = 0$$

is satisfied.

(This follows essentially from (3.8).) The theorem is interesting because the geometrical property stated means that the vortex-lines consist of the same fluid elements during all the time. The condition is satisfied, e.g., if the streamlines are "conformal geodesics", i.e. if they are geodesics with respect to a metric conformally related to the natural one, g_{ab}. But in this case a stronger theorem holds:

<u>Theorem 3.2.</u> In a conformally geodesic flow (with $\dot{u}_a = -h_a^{\ b}\cdot(\log w)_{,b}$) the propagation law

(3.11) $$h^a_{\ b}(wl^2\omega^b)^{\cdot} = \sigma^a_{\ b}wl^2\omega^b$$

holds. (l is the scale-factor defined in (3.7).)

Essentially from (3.8) one can deduce a kinematical characterisation of "isometric" flows (=group motions, cf. sec. 2,2) :

Theorem 3.3. A rigid motion is a group motion if and only if the vortex vector is FERMI-propagated along the streamlines and the acceleration vector points always to the same neighbouring particles:

(3.12) $\Theta = \sigma = 0$: $\dot{u}_{[a,b]} = 0 \Leftrightarrow \left\{ h^a_b \dot{\omega}^b = 0,\ h^a_b \ddot{u}^b = \omega^a{}_b \dot{u}^b \right\}$

This statement admits the following intuitive interpretation: A gravitational field is stationary if and only if it is possible for a suitably chosen cloud of particles to move rigidly and with constant angular velocity under the influence of internal (non-gravitational) forces. If, moreover, such a motion exists which is in addition irrotational, the field is static. (The particles need not be test particles because no field equations are assumed.)

A theorem useful for the discussion of solutions of the field equations corresponding to rigid body-motion (Ehlers 1957, 1959; Rayner 1959) is this:

Theorem 3.4 The trajectories of a group motion are RICCI - lines if and only if a scalar W exists with

(3.13) $\omega_a = e^{-2U} W_{,a}$

(U is the acceleration-potential from eq. (2.1)). If this is true W satisfies

(3.14) $(e^{-4U} W^{,a})_{;a} = 0$.

Finally we formulate a consequence of equs. (3.3), (3.5) which is analogous to certain theorems in the theory of gravitational pure radiation fields (see, e.g., Jordan-Ehlers-Sachs 1961), but is independent of field equations:

<u>Theorem 3.5</u> Let K be a shearfree, normal, timelike congruence in a space-time W. If K is even rigid or if K consists of RICCI-lines, then the conformal curvature tensor of W has PETROV-type I with real eigenvalues, and u^a is an eigenvector of R_{ab} and a principal vector of C_{abcd}.

This theorem is a slight generalization of a well-known theorem about static space-times.

§ 4. Remarks on the dynamics of ideal fluids and incoherent matter.

The fundamental dynamical law

$$G_{ab} - g_{ab} = - T_{ab} \tag{4.1}$$

of EINSTEIN's theory of gravitation is a general scheme which gets a well-defined physical content only if the source-term T_{ab} is specified according to the type of matter considered.

If we assume only that a mean-velocity u^a is defined besides T_{ab}, we may apply the kinematics developed above to the congruence defined by u^a.

Further, $\mu = T_{ab} u^a u^b$ is then defined to be the mean energy density, and $p = \frac{1}{3}(T + \mu)$ - with $T = T^a{}_a$ - is the mean pressure.

If these definitions and (4.1) are combined with equs. (3.4) and (3.7) we obtain

(4.2) $$3\frac{\ddot{l}}{l} + 2(\sigma^2 - \omega^2) - \dot{u}^a{}_{;a} + \frac{1}{2}(\mu + 3p) - \Lambda = 0.$$

Consequently, the quantity $-\Lambda + \frac{1}{2}(\mu + 3p) - 2\omega^2$ is the source of the acceleration field in rigid flows. This fact can be considered to be a relativistic analogue of GAUSS's law in ordinary gravitational theory, if the latter is formulated with respect to a rigidly rotating frame of reference. In an isometric flow, (2.1) leads to the "POISSON"-equation

(4.3) $$U'^a{}_{;a} = -\Lambda + \frac{1}{2}(\mu + 3p) - 2\omega^2 \quad (\dot{U} = 0).$$

If this equation is transformed into a three-dimensional divergence-relation it turns out that $(\mu + 3p)e^U$ should be considered to be the effective density of (active) gravitational mass. (This should be compared with WHITTAKER 1935 and PIRANI 1956.)

If one considers ideal fluids with a pressure-density relation the "conservation law"

(4.4) $$T^{ab}{}_{;b} = 0$$

is known to imply that the congruence of the streamlines is

conformally geodesic. Consequently, the vortex-theorems stated above are applicable. It should also be noticed that theorem 3.5 applies: A shearfree, irrotational, isentropic flow of a perfect fluid is possible only in a gravitational field the conformal curvature tensor of which has PETROV-type I and real eigenvalues. This corresponds to the idea that such fields do not contain gravitational radiation; for such radiation would presumably cause shear motions of the matter interacting with the field.

The simplest model of matter is that of dust, characterized by

$$T_{ab} = \varrho\, u_a u_b \,. \tag{4.5}$$

To the (geodesic) congruence of such matter the kinematical theorems are applicable, leading, e.g., to the statement that in a shear-free flow $l^2\omega^a$ is parallely propagated along the streamlines, that conformal, expanding motions are irrotational, and that rigid motions are isometric, with a harmonic vortex vector ω^a ($\omega_{[a,b]} = \omega^a{}_{;a} = 0$) .

With respect to incoherent matter (4.5) we state the following

<u>Theorem 4.1.</u> The field equations (4.1) with the source (4.5) are equivalent to the following set of equations:

$$h^a_b(\omega^{bc}{}_{;c} - \sigma^{bc}{}_{;c} + \tfrac{2}{3}\Theta^{,b}) = 0, \tag{4.6}$$

$$3\,\frac{\ddot{l}}{l} + 2(\sigma^2 - \omega^2) + \tfrac{1}{2}\varrho - \Lambda = 0\,, \tag{4.7}$$

$$- h_{ab} R^{bc} h_{cd} = (\Lambda + \tfrac{1}{2}\varrho)\, h_{ad} \tag{4.8}$$

Corrollary 1. If $\sigma = 0$ (isotropic expansion), (4.7) can be replaced by the "energy conservation law"

(4.9) $$3\dot{l}^2 + \frac{2\Omega^2}{l^2} - \frac{M}{l} - \Lambda l^2 \equiv E, \quad \dot{E} = 0$$

with

(4.10) $$\omega l^2 \equiv \Omega, \quad \dot{\Omega} = 0 \; ; \quad \rho l^3 \equiv M, \quad \dot{M} = 0$$

Corrollary 2. In an irrotational motion, (4.8) can be replaced by the geometrical statement:

(4.11) $$K(x^a, e^a) = \frac{1}{3}(\Lambda + \rho + \sigma^2 - \frac{1}{3}\theta^2) + l^{-3}(l^3\sigma_{ab})^{\cdot} e^a e^b$$

is the GAUSS-curvature of that surface F generated within the hypersurface R (R containing x^a and orthogonal to the streamlines) by the geodesics (with respect to the metric induced in R) starting from x^a in directions orthogonal to e^a ($e^a u_a = 0$).

The first corrollary is the covariant formulation of RAYCHANDHURI's equation (Raychandhuri 1955) which is important to cosmology. The second corrollary illustrates the manner in which the distribution and relative motion of matter is related to the spatial curvature.

Both corrollaries together may be used to construct and characterize in a simple and intuitive way the FRIEDMANN-LEMAITRE-universes.

We also see from (4.11): In an irrotational flow of incoherent matter the rest spaces have constant curvatures if and only if the tensor $l^3\sigma_{ab}$ is parallely propagated along the streamlines.

By this remark and the above theorem it is possible to construct simple generalizations of the FRIEDMANN-models; examples are some models given by SCHÜCKING and HECKMANN at the Solray-conference 1958. (These authors used a different method.)

§ 5. Kinetic theory of gases. H-theorem. Equilibrium distributiones in a gravitational fields.

We now consider matter not as continuously distributed, but to be a statistical ensemble of particles of rest mass m interacting only by elastic collisions. We assume space-time to be curved, but smooth enough such that "physically infinitesimal" space-time regions of the type required to define densities and momenta of the distribution function and to formulate the BOLTZMANN equation can be considered as flat. The metric field g_{ab} must accordingly be considered as the macro-field (to use a term from the theory of electrons).

To define the (one-particle-) distribution function $F(x,p)$ let x be an event, dx^{abc} a space-like 3-cell at x with dual $d\overset{*}{x}_a$, and dp^{abc} a 3-cell at p^a in momentum-space contained in the hyperboloid

(5.1) $$p^a p_a = -m^2 .$$

Then, by definition,

(5.2) $$F(x,p)\,|d\overset{*}{x}_a\, d\overset{*}{p}{}^a|$$

is the number of particles whose worldlines intersect dx^{abc} with momenta contained in dp^{abc}. Obviously F is a scalar.

F satisfies the BOLTZMANN equation

(5.3) $$p^a F_{,a} = \iiint (F'' F''' - FF')\, W(p, p'; p'', p''')\, dP' dP'' dP'''$$

formulated by SASAKI in 1958. The partial derivative $F_{,a}$ is to be understood with respect to x^a, p^a being parallely propagated. The first argument in the F's is always x, the second p'', p''', p, p', respectively. Moreover we have put

(5.4) $$d\overset{a}{p}_a \equiv p_a \, dP .$$

$W(p, p'; p'', p''')$ describes the probability of collisions $p, p' \to p'', p'''$. We assume, as usual,

(5.5) $$W(p, p'; p'', p''') = W(p', p; p''', p'') = W(p'', p'''; p, p').$$

The first moment

(5.6) $$g^a \equiv \int g(p)\, p^a\, F(x, p)\, dP$$

of an arbitrary function $g(p)$ satisfies, in consequence of (5.3) and (5.5), the transport-equation

(5.7) $$g^a{}_{;a} = \frac{1}{4} \iiiint (g + g' - g'' - g''')(F'' F''' - F F')\, W dP dP' dP'' dP'''$$

which also holds for tensors $g^{a\ldots}$ instead of g.

If g is an additive collision invariant it follows that

$g^a{}_{;a} = 0$. This implies:

1. The matter-flux density

(5.8) $$\rho u^a \equiv \int m p^a F\, dP \qquad (u_a u^a = -1)$$

satisfies the equation of continuity

(5.9) $$(\rho u^a)_{;a} = 0.$$

ρ is the density of proper mass with respect to the local rest frame of the fluid defined by u^a, the barycentric mean velocity.

2. The energy-momentum tensor

(5.10) $$T^{ab} \equiv \int p^a p^b F\, dP$$

has vanishing divergence:

(5.11) $$T^{ab}{}_{;b} = 0.$$

Next we define the entropy flux density S^a by

(5.12) $$S^a \equiv - \int p^a F \log F\, dP.$$

We obtain, differentiating,

$$S^a{}_{;a} = - \int (1 + \log F)\, F_{,a}\, p^a\, dP.$$

The same transformation which led to eq. (5.7) can now be applied; the result is

(5.13) $$S^a{}_{;a} = \frac{1}{4} \iiiint \log \left(\frac{F'' F'''}{F\ F'} \right) (F'' F''' - F F') W dP dP' dP'' dP'''.$$

Therefore the following H-theorem is valid:

<u>Theorem 5.1</u> The entropy production density

$$S^a{}_{;a} \equiv G$$

is non-negative and vanishes only in a space-time region R if everywhere in R $\log F$ is an additive invariant under collisions.

We now make the assumption suggested strongly by the laws of relativistic collisions that - corresponding to H. GRAD's well-known theorem in the non relativistic theory of elastic collisions - the general collision invariant is given by $c + \xi_a p^a$. Then the preceding theorem shows: The equilibrium-distributions of a relativistic gas have the form

$$F(x,p) = C(x)\, e^{\xi_a p^a} \quad . \tag{5.14}$$

From this formula and (5.8,10) we obtain

<u>Theorem 5.2.</u> The local equilibrium distribution of an ideal gas is given by

$$F(x,p) = \frac{\rho(x)\,\xi}{4\pi m^4 K_2(m\xi)}\, e^{\xi u_a p^a} \quad . \tag{5.15}$$

The corresponding energy tensor is that of an ideal fluid with the equation of state

$$\mu + p = \rho\, G\left(\frac{\rho}{p}\right) \tag{5.16}$$

in which

$$G(y) = \frac{2}{y} - \frac{K_2'(y)}{K_2(y)}\,, \quad K_2(y) = \int_0^{\infty} e^{-y\cosh z} \cosh(2z)\, dz. \tag{5.17}$$

The distribution-parameter ξ is related to the pressure p and the proper density of proper mass, ρ, by

$$m\,\xi = \frac{\rho}{p}\,. \tag{5.18}$$

(This theorem has been established by SYNGE 1958; the difference between his treatment and the present one consists in that SYNGE uses the BOLTZMANN-method of the most probable distribution as derived from permutabilities, whereas we have used the collision equation and have thus established what has been called the "dynamical" proof of the equilibrium-states.)

We now ask: What conditions are imposed on the "state fields" $\rho(x)$, $\xi(x)$ and $u^a(x)$ by the collision equation (5.3)? The answer is given by

<u>Theorem 5.3.</u> An equilibrium distribution (5.15) is possible only in a stationary gravitational field . More specifically, $\xi^a = \xi u^a$ has to be a KILLING vector, and consequently the flow is a group motion. Moreover,

$$\frac{p\,\xi^2}{K_2(m\xi)} \equiv A \tag{5.19}$$

must be constant in the domain of the flow. These conditions are also sufficient.

The proof of this theorem is straightforward.

Comparing this result with the classical one obtained by BOLTZMANN we observe: The reversible flows of an ideal gas in a gravitational field are kinematically more restricted in the relativistic than in the classical treatment; in the relativistic case they are not only shear-free but even rigid. This implies:

In the isotropically expanding universes of relativistic cosmology the substratum can - in contrast to the Newtonian analogue - not be thought of as being an ideal gas in local equilibrium. (This gives a partial answer to a question put forward by HECKMANN , see Heckmann 1942.)

Now we decompose the entropy-flux density into a convection and a conduction part:

(5.20) $$S^a = \varrho s u^a + s^a \quad (u_a s^a = 0)$$

In an equilibrium distribution s^a vanishes according to equs. (5.12) and (5.14), and s , the specific entropy, is seen - after a little calculation - to be related to the specific energy $u = \frac{\mu}{\varrho}$ by $\xi(du + pdv) = ds$. This shows that $\frac{1}{\xi} \equiv T$ is the temperature of the gas, and therefore we may infer from theorem 5.3 , remembering the definition of the scalar gravitational potential U in stationary fields:

Theorem 5.4. In an ideal gas, being in thermodynamic equilibrium in a gravitational field, the temperature and the gravitational potential are related by

(5.24) $$T e^U = \text{const.}$$

This law has been established by R.C. TOLMAN who considered (phenomenologically) the equilibrium of isotropic electromagnetic radiation in a static gravitational field. It is a confirmation of the inner consistency of relativity that a quite different

model and a different method again lead to this law. (It should be noticed that the "gravitational" potential might be a centrifugal potential also.)

Finally we remark that this treatment of the relativistic gas has been stimulated by the works of SYNGE 1958, TAUB 1948 and SASAKI 1958, and should be considered as a completion of these former investigations.

References

Ehlers, J., 1957, Dissertation Univ. Hamburg,

" " 1959, Referat auf der Konferenz über Allgemeine Relativitätstheorie in Royaumont, Juni 1959

Eisenhart, L.P., 1924, Trans.Amer.Math. Loc.26, 205

Jordan, P., J. Ehlers und R.K. Sachs, 1961, Akad.Wiss.Mainz Abh. Math.-Nat.Kl. 1961 Nr. 1

Lichnerovicz, A., 1955, Theories relativistes de la gravitation et de l'electromagnetisme, Paris

Pirani, F.A.E., 1956, Acta Phys. Polon. 15, 389

Raychandhuri, A., 1955, Phys. Rev. 98, 1123

Rayner, C.B., 1959, C.R. Paris 248, 929; 2725; 249, 1327

Salzman, G., and A.H. Taub, 1954, Phys. Rev. 95, 1659

Sasaki, M., 1958, Planck-Festschrift, S. 129

Schücking, E., und O. Heckmann, 1958, Ber. Solray Konf. Brüssel 1958

Synge, J.L., 1937, Proc. Lond. Math. Soc. 43,

" " 1958, The Relativistic Gas, Amsterdam

" " 1960, Relativity: The General Theory, Amsterdam

Taub, A.H. 1948, Phys. Rev. 74, 328

Tolman, R.C., 1934, Relativity, Thermodynamics and Cosmology, Oxford

Witten, L. (Editor), 1961, The Theory of Gravitation, Wiley; ch.2 by Ehlers, J., and W. Kundt

Whittaker, E.T., 1935, Proc. Roy.Soc. 149, 384

Fotodruck Mikrokopie G.m.b.H München 2, Weinstr. 4